TASTE AND ODOR EVENT DYANMICS OF A MIDWESTERN FRESHWATER RESERVOIR

Chase Steven Howard

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Master of Science
in the Department of Earth Sciences,
Indiana University

November 2020

Accepted by the Graduate Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Master of Science.

Master's Thesis Committee

_____
Gregory K. Druschel, PhD, Chair

_____
Pierre-André Jacinthe, PhD

_____
Christine J. Picard, PhD

DEDICATION

Dedicated to my family, which is everything to me. My parents, Steve and Amy Howard, who have given me immeasurable emotional and financial support to shape me into the man I am today. My fiancé and future wife, Lauren Hucek, who is the most supportive and understanding woman that I know. And my pets, Henry, Dobby, and Slippy, who are always there during late nights of work at the computer.

ACKNOWLEDGEMENT

The contributions of numerous people, including advisors, colleagues, friends, and family have pushed this work forward to completion. I am extremely grateful for the support I have received during my years at the IUPUI Earth Sciences department, without which this endeavor would not have been possible.

I want to thank Dr. Gregory Druschel for being my advisor and mentor during both my undergraduate and graduate studies. Working in his lab has afforded me many opportunities I would not have found elsewhere, all culminating in me becoming a better scientist. I have always been deeply humbled by the amount of responsibility Greg has entrusted me with in allowing me to lead a portion of the lab's work at Eagle Creek Reservoir. The independence with which Greg allowed me to work, while still always being available for guidance and help, has imparted on me a profound sense of confidence and accomplishment. I could not have asked for a better advisor.

My committee members, Dr. Pierre-André Jacinthe and Dr. Christine Picard, deserve praise as well. Dr. Jacinthe taught me the fundamentals of soil science during my coursework and gave important input that helped guide the work for my thesis. Dr. Picard and her student Ioanna Koltsidou developed a set of targeted genetic techniques for our work at the reservoir and provided us with important data used in my thesis. Additionally, Dr. Picard gave important feedback regarding data analysis and was responsible for the inclusion of Boosted Regression Trees in my thesis.

Gathering samples and generating data was a group effort spanning several years, and I would like to thank all of those members of the team. First, the other graduate students Igor Ogashawara and Ioanna Koltsidou who were largely responsible for the

spectral and genetic components of the data, respectively. They spent the better part of three seasons working side by side with me on a boat, and I am deeply thankful for their help. Likewise, I would like to thank Sarah Denny, Tyler Lawson, Garrett Scott, Lex Cantrell, and Natalie Nichols, all undergraduates who assisted in sampling and lab work throughout the project. Last but not least are the project's drone pilot Jeremy Webber III and his assistant Lynn Thomas.

This project was made possible by Citizens Energy Group (CEG) Grant 123752, titled "Taste and Odor compounds in Eagle Creek Reservoir: Developing spatial and temporal tools for early identification of antecedent conditions and problem microbial blooms" along with the support of CEG's White River Water Treatment Plant labs. I would like to thank CEG and specifically our lab's contacts at the company, Mark Gray and Leo Sparks, for their financial support and the analyses their lab assisted in.

Additionally, a great deal of logistical support for the project was provided by IUPUI's Center for Earth and Environmental Science (CEES) under the direction of Dr. Victoria Schmalhofer. CEES gave the lab access to both the boat used for sampling and to undergraduate assistants who could help during field work.

My time at IUPUI was greatly improved by my interactions with the Earth Sciences department and I would like to express my gratitude to a number of faculty members. These include Dr. Gabriel Filippelli, Dr. William Gilhooly III, Dr. Lixin Wang, and Dr. Kathy Licht, all of whom instructed me in various fields. In addition, Cathy Chouinard and Cheryl Montgomery who maintain the department's day to day functions and oversee lab logistics.

Finally, I would like to thank members of the Druschel lab group (many of whom have already been mentioned) for their input, suggestions, and support. John Shukle in particular was an invaluable resource when it came to statistical analysis, computer coding, and interpreting results. He was fundamental in my decision to learn the R programming language. Others in the group who have provided comments and discussion throughout the years but have yet to be recognized include Martin Kurek, Dr. Rick Whitman, and Patrick Cavanagh.

Chase Steven Howard

TASTE AND ODOR EVENT DYANMICS OF A MIDWESTERN FRESHWATER RESERVOIR

Eagle Creek Reservoir (ECR), located in the Midwestern U.S., is a freshwater limnic system plagued by seasonal Harmful Algal Blooms (HABs) which generate water-fouling Geosmin (GSM) and 2-Methylisoborneol (MIB) Taste and Odor (T&O) compounds. Past investigations of T&O event dynamics have identified *Actinomycetes* as responsible for MIB production and several genera of cyanobacteria for GSM production. During 2018, a temporally and spatially expansive sampling regimen of the reservoir was carried out and a battery of biological, chemical, physical, and hyperspectral experiments performed. The resulting data was analyzed using time series, cross-correlation, lag time, and multivariate analyses as well as machine learning algorithms to pick apart and interrogate any relationships between HABs, T&O events, and environmental parameters. The results show that local weather and watershed conditions exert significant control over the state of the reservoir and the behavior of the algal community. GSM and MIB peaked during early May under well-mixed, cold, and nutrient-rich water column conditions, then declined under summer thermal stratification before making a small resurgence during late season mixing. Bloom die-off and decay was effectively ruled out as a mechanism controlling T&O concentrations, and no links were found between T&O concentrations and algal biomass. Strong evidence was found that GSM/MIB concentrations were a response by bloom microbes to changing nutrient conditions within the reservoir, and it was determined that nutrient fluxes from the watershed 30-40 days prior to peak T&O concentrations are likely instrumental in the development of the slow-

growing microbes characteristic of the reservoir. Attempts were made to assess spatial and temporal variability but no significant spatial differences were identified; differences between sampling sites were far smaller than differences between different sampling dates. The findings here add to the growing body of literature showing T&O and HAB dynamics are more closely linked to the relative abundance and speciation of nutrients than other parameters. Additionally, these findings carry important implications for the management of ECR and other similar freshwater reservoirs while highlighting the importance of reducing watershed eutrophication.

Gregory K. Druschel, PhD, Chair

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BRT	Boosted regression tree
CEES	Center for Earth and Environmental Science
CEG	Citizens Energy Group
Chl-a	Chlorophyll-*a*
DGage	Variable relies on data taken at the dam gage (USGS 03353451)
DN	Dissolved nitrate
ECR	Eagle Creek Reservoir
GSM	Geosmin
HAB	Harmful algal bloom
MIB	2-Methylisoborneol
N	Nitrogen
NMDS	Non-metric multi-dimensional scaling
P	Phosphorus
PC	Phycocyanin
PCA	Principal component analysis
qPCR	Quantitative polymerase chain reaction
SGage	Variable relies on data taken at the dam gage (USGS 03353200)
SOP	Standard operating procedure
SRP	Soluble reactive phosphorus
T&O	Taste and odor
TDP	Total dissolved phosphorus
TN	Total nitrogen

TP	Total phosphorus
TSS	Total suspended solids
USGS	United States Geological Survey
VOC	Volatile organic compound

Chapter 1 – A review of relevant literature and an overview of the 2018 ECR study

Introduction:

HAB prevalence and environmental impact

Harmful Algal Blooms (HABs) result when the overgrowth of algae, a catch-all term encompassing a plethora of both photo- and heterotrophic microorganisms in aquatic environments [Ramsing *et al.* 1996, Heisler *et al.* 2008], becomes substantial enough to negatively impact overall water quality or the local ecosystem. HABs are a pronounced and growing problem for most freshwater bodies around the globe and have garnered an increasing amount of attention from the scientific community in recent decades [Smayda 1997, Juttner and Watson 2007, Heisler *et al.* 2008, Anderson 2014]. The rise in the frequency and intensity of HABs is thought to be the result of degraded water quality from increased nutrient pollution, referred to as eutrophication [Paerl *et al.* 2001, Heisler *et al.* 2008], in combination with increasing temperatures driven by global warming/climate change [Shatwell *et al.* 2008, Paerl and Huisman 2009]. As such, the current trend of worsening HAB events is expected to continue for many locations around the world [Wells *et al.* 2020].

HABs can pose a threat to the local environment in several ways, perhaps most notably by causing water column anoxia. Heterotrophic organisms within sufficiently dense blooms are capable of depleting the available oxygen and creating anoxic conditions, which in turn causes benthic mortalities and/or fish kills [Smith *et al.* 1999, Diaz 2001, Anderson *et al.* 2002]. Aside from severe oxygen depletion, HABs have been found to cause significant ecological changes which may affect the overall ecosystem

[Miguéns and Valério 2015]. In some localities, HABs are hazardous to human health as they can be responsible for producing a wide range of neuro- and hepatotoxins [Carmichael 2001, Hamill 2001]. The adverse outcomes of HABs necessitate mitigation and remediation, especially in freshwater bodies used for drinking water, leading to substantial economic costs for governments and businesses [Dodds *et al.* 2008].

Role of nutrients and eutrophication

The growth and proliferation of HABs has been studied extensively and it is generally established that nitrogen (N) and phosphorus (P) are the limiting nutrients for phytoplankton growth in freshwater lakes [Lean *et al.* 1987, Xu *et al.* 2010]. The importance of each nutrient can vary from system to system, as several microorganisms commonly found in blooms have demonstrated N-fixation capabilities [Leschine *et al.* 1988, Fay 1992]. The molecular form of both nutrients and their respective sources is crucial to understanding the dynamics of nutrient-bloom interactions for an individual freshwater body. N and P can be derived from any number of natural or anthropogenic sources, leading to external and/or internal nutrient loading in a system [Xu *et al.* 2010]. Human activity within watersheds has greatly increased the rate/amount of nutrients transported downstream, and in general increased urbanization is directly correlated with increased nutrient pollution [Valiela *et al.* 1992]. Point sources in watersheds include mine drainage/run-off and sewage outflows though these are often less important than non-point sources, which include fertilizer/manure application/run-off and fossil fuel combustion [Anderson *et al.* 2002]. The over-application of fertilizer and manure to fields, along with high density animal husbandry, tend to enrich soils in nutrients which can then be transported through leaching or soil erosion [Carpenter *et al.* 1998]. A firm

grasp of these nutrient dynamics as they relate to a particular system is often key to knowing the likelihood and timing of HABs.

The impact human activity on the global N cycle has had is substantial and nitrogen's availability and mobility have been greatly enhanced through anthropogenic modification [Vitousek *et al.* 1997]. A majority of nitrogen is contained within a massive and virtually unavailable pool of nitrogen gas (N_2) in the Earth's atmosphere. A small portion of N_2 is converted through biologically mediated processes to more bioavailable chemical forms [Delwiche 1970]. This crucial step in the cycle, termed N fixation, converts N_2 to ammonia which can then undergo further conversion through nitrification to form nitrites and nitrates. Human activity has effectively doubled the speed of this transfer of N from the atmospheric pool to land and water, by industrial N fixation (fertilizer creation and application), fossil fuel combustion (production of nitrogen oxides), extensive sowing of N-fixing crops (leguminous and forages), and the burning/clearing of biomass [Vitousek *et al.* 1997]. This accelerated N fixation drives eutrophication of waters within a catchment or watershed; a British study using ^{15}N isotope spiked fertilizer showed that over 90% of the added ^{15}N could be found within algal biomass during peak blooms [Stewart *et al.* 1982].

In many aquatic systems, focus is often placed on P sources, dynamics, and cycling due to its relative scarcity compared to N. P can be found in the land or water as part of a complex cycle which is generally less studied than the N cycle [Smil 2000]. External inputs of P into a body of water stem from a number of point and non-point sources [Anderson *et al.* 2002] like N, but lake sediments can serve as a large and dynamic reservoir for P leading to internal loading of the nutrient in some systems

[Søndergaard *et al.* 2003]. The speciation, form, and availability of this sediment-bound P is highly dependent on both the chemical composition of the water column and the sediment itself [Petterson 1998]. Hydrous oxides of manganese and iron, reactive particulate carbon, hydrous oxides of aluminum and silicon, and clay minerals have long been known to sorb metals and ions in soils and freshwater sediments, though the first two categories exert principal control [Jenne 1968]. Sorption occurs across mineral surfaces at binding sites containing functional groups and a surface charge develops once this solid surface is in contact with an aqueous phase [Smith 1999]. These functional groups are often ionizable, leading to variable surface charges that change with pH and redox potential. Changes to surface charge are often thought of in the context of the point of zero charge (pzc), the pH at which the net charge of the mineral's surface is zero. Below this pH, positive surface charges are generated and above this pH negative charges result, attracting absorption of anions and cations, respectively. Inorganic and charged species of phosphorus easily sorb to these surfaces if/when opposite charges are present, but phosphorus can also be mineralized as primary P minerals like apatite or secondary P minerals of Ca, Fe, and Al phosphates which represent a significant P pool in some sediments [Pierzynski *et al.* 2005, Łukawska-Matuszewska *et al.* 2013]. The fraction of inorganic P bound to iron/manganese hydroxides experiences sorption under favorable pH/redox conditions, but changes in pH affect the capacity of the mineral to sorb orthophosphate while sufficiently reducing conditions lead to the reduction of Fe(III) to Fe(II) and subsequent dissolution of the hydroxide mineral [Lijklema 1980]. This dissolution releases sorbed P and any P that had been incorporated into the mineral's structure, resulting in fluxes of P into the overlying water column [Pettersson 1998,

Christophoros and Fytianos 2006, Smith *et al.* 2011]. The process of P introduction to a freshwater lake, its sedimentation/mineralization/immobilization in the system, and seasonal hypoxia and/or thermal stratification causing release can mobilize P in fluxes several times greater than externally loaded phosphorus concentrations [Søndergaard *et al.* 2013].

HAB distributions

The spatio-temporal distribution of HABs is not as well studied as nutrient importance in phytoplankton growth, and the role of nutrient concentrations versus other environmental factors is still highly debated [Flynn 2010, Davidson *et al.* 2012, Glibert *et al.* 2013, Anderson 2014]. Several studies have shown that local hydrology [Šimek *et al.* 2011] and water column mixing or stratification [Ramsing *et al.* 1996] may exert significant control over the spatial distribution of HABs and their timing. It has been well established, however, that the spatial distribution of planktonic organisms is highly heterogeneous [Pinel-Alloul 1995]. The phenomena is referred to as "planktonic patchiness" and is important when assessing the effects of HABs on overall water quality and ecosystem health. Planktonic patchiness occurs across a number of scales spanning from the microscopic to the size of oceans and is a function of both abiotic and biotic processes [Pinel-Alloul & Ghadouani 2007]. On smaller scales, biotic processes play the most important roles, while physical abiotic processes dominate with increasing scale. Pinel-Alloul & Ghadouani (2007) posit that these scales should be thought of in a hierarchical manner, that is, patchiness at smaller scales contributes to the structure of patchiness at larger scales. Studies are often constrained to a single scale with coarse sampling, for example a handful of sampling sites spread every few hundred meters

across a reservoir as opposed to sites located every few meters, which works for many applications but makes studying HAB heterogeneity difficult. Physio-chemical conditions in the water column can also mirror this heterogeneity, with significant spatial differences in nutrient measurements and redox conditions [Barker *et al.* 2010] possibly driving HAB distribution.

T&O compounds in the environment

One common process occurring with some HAB events are the production of secondary metabolites, including a class of Volatile Organic Compounds (VOCs) referred to as Taste and Odor (T&O) compounds [Juttner and Watson 2007, Krishnani *et al.* 2008]. These molecules are easily volatilized into the gaseous form, and human senses can detect them even in exceedingly low concentrations on the range of 10-20 ng/L [Ömür-Özbek *et al.* 2007]. This is a problem especially for drinking water as these T&O molecules have an easily observable negative impact on overall water quality [Srinivasan and Sorial 2011] which can cause significant economic damage. It should be noted that T&O compounds are considered to be non-toxic and largely unrelated to neurotoxin production in limnic systems, at least for the Midwestern U.S. [Graham *et al.* 2010]. T&O events can be remediated through a battery of water treatment techniques (typically absorption to granular activated carbon (GAC) or powdered activated carbon (PAC), advanced oxidative processes using UV/ozone/H_2O_2, or biologically activated filter mediums) [Srinivasan and Sorial 2011] or through algaecide application [Pascual and Tedesco 2004], but with limited and varying efficacy often tied to the speciation of microorganisms within the blooms [Bishop *et al.* 2014, Kansole and Lin 2017, Shen *et al.* 2019]. Of the T&O compounds, Geosmin (GSM) and 2-Methylisoborneol (MIB), are the

main culprits in surface freshwaters affected by HAB T&O events [Lanciotti *et al.* 2003, Ma *et al.* 2013].

T&O producers and T&O events

GSM and MIB are produced by microorganisms mostly within Actinobacteria [Gerber and Lechevalier 1965, Juttner 1990] and Cyanobacteria phyla [Juttner and Watson 2007, Li *et al.* 2007, Su *et al.* 2015]. Species and their genetic sequences responsible for T&O production have been identified [Bentley and Meganathan 1988, Komatsu *et al.* 2008, Watson *et al.* 2016] and the biosynthesis of both molecules is produced by three main pathways. These are the 2-methylerythritol-4-phosphate (MEP), mevalonate (MVA), and leucine pathways, with MEP as the major biosynthetic route for many bacterial species and the leucine biosynthesis pathway is thought to only play a minor role [Juttner and Watson 2007]. The MEP pathway has been observed in *Streptomyces* bacteria (Actinobacteria) and the necessary genes are found in several cyanobacteria species as well. The MVA pathway has been observed in *Streptomyces* too, with work by Bentley and Meganathan (1981) suggesting it was the predominant pathway for GSM production. Later evidence did not support MVA over MEP as the dominant mechanism, but showed MVA genetics were also present in some cyanobacteria. The enzyme involved in the MVA pathway also serves other metabolic functions, however, so its presence did not necessarily indicate involvement in GSM synthesis [Juttner and Watson 2007]. The purpose of T&O molecule production is not fully understood, but they are considered secondary metabolites and therefore unlikely to influence the growth or reproduction of their producers. It has been suggested that these chemicals influence microbial community dynamics and serve as either infochemicals

used for communication or as antimicrobial agents used to fight competitors [Watson *et al.* 2003, Tyc *et al.* 2017].

Remediation through algaecide application has disparate outcomes based on bloom speciation and community structure [Bishop *et al.* 2014, Kansole and Lin 2017, Shen *et al.* 2019]. For example, repeated observations at Eagle Creek Reservoir (ECR) have demonstrated that cyanobacteria are far more susceptible to algaecides than actinobacteria [Pascual and Tedesco 2004, Clercin and Druschel 2019]. Most algaecides are copper-based, including those used at ECR, and the differences in response to algaecide treatment between cyanobacteria and actinobacteria communities are likely due to each microorganism's ability to internally regulate copper concentrations. Copper algaecide primarily functions by elevating copper concentrations to a cytotoxic level within the cell, where the element inhibits photosystem II activity and electron transport systems of cyanobacteria [Shen *et al.* 2019]. This leads to eventual cell lysis and death. Certain actinobacteria like *Streptomyces* have demonstrated a profound efficiency at eliminating copper from themselves and were found to possess cupric reductase enzymes [Albarracín *et al.* 2008]. In the infectious disease literature, cupric/ferric reductase has been found in multiple bacteria dependent on iron for growth [Nyhus and Jacobson 1999]. Actinobacterial resistance to copper-based algaecides is probably linked to this ability to regulate copper transport across the cell membrane, while cyanobacterial susceptibility stems from their lack of such a mechanism.

The required circumstances for T&O compounds to be upregulated and expressed are not well understood. HABs and the T&O events they can generate are often a seasonal occurrence [Ramsing *et al.* 1996, Li *et al.* 2007, Su *et al.* 2015] and the two

chemicals can be found both during HAB events and afterwards, being released with the decay of the bloom's microorganisms [Ma *et al.* 2013]. Past work pointed to overall bloom biomass and development controlling GSM concentrations rather than changing biosynthesis rates in response to environmental conditions [Wnorowski 1992]. However, lab experiments showed that GSM and MIB production rates were linked to biomass for certain *Streptomyces* species but not others [Podduturi *et al.* 2018]. Furthermore, a comprehensive multi-year study of the Kansas River found no relationships between GSM/MIB and algal biomass, cyanobacterial abundance, or actinomycetes abundance [Graham *et al.* 2018]. Recent success has been found in analyzing the ratios of available nutrients, with the principal that this ratio is an indicator of N-limited (N:P is low) and P-limited (N:P is high) states. Across Midwestern reservoirs, low TN:TP ratios were seen to favor cyanobacterial dominance and growth while low NO_3:NH_3 ratios were associated with increased T&O production [Harris et al 2016]. Other studies have found similar associations between low NO_3:NH_4 ratios and heightened T&O production [Perkins *et al.* 2019]. One recent study of Acton Lake in southwestern OH, U.S. even demonstrated that variability in nitrate, ammonium, and phosphorus ratios was strongly correlated to seasonal algal community trends and shifts in the state of nutrient limitation [Andersen *et al.* 2020].

Study site:

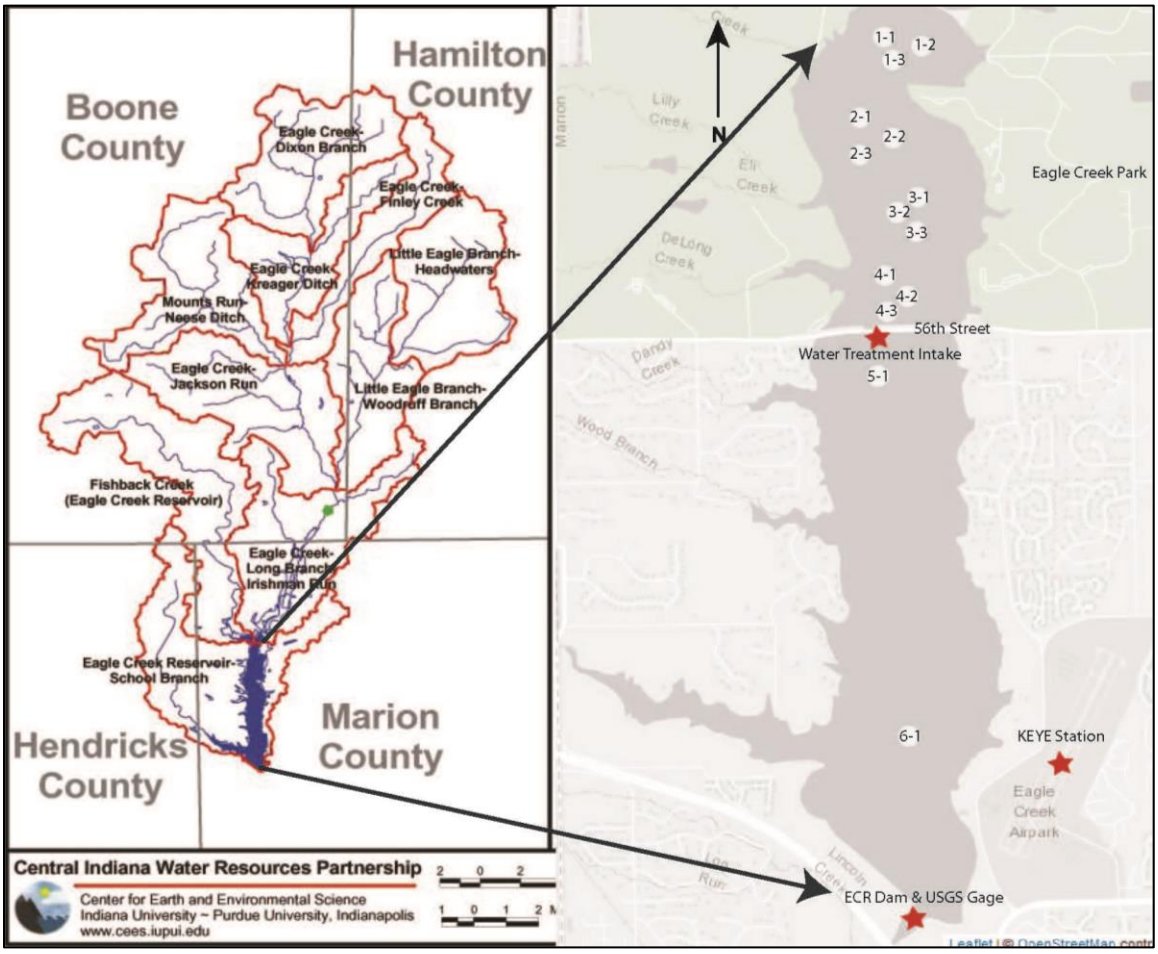

Figure 1.1: Overview of Eagle Creek Reservoir (ECR). Left pane modified from Tedesco *et al.* (2005) and illustrates the extent of ECR and its watershed. Right pane displays the reservoir itself, with sampling locations ("X-Y") and key locations (red stars) marked.

Overview of ECR

Eagle Creek Reservoir (39˚51'20"N, 86˚17'39"W), located in central Indiana, is an important freshwater resource for the city of Indianapolis that provides drinking water for over 80,000 citizens and serves as a recreational area for local residents [Pascual and Tedesco 2004]. The reservoir was constructed in 1969 as a means of flood control for the western portion of the city and later became a drinking water source as well. The

watershed (HUC 05120201120) is substantially larger than the surface area of the reservoir at roughly 420 km^2 compared to 5.7 km^2. The land surface within the watershed is dominated by agriculture at almost 60% of the total area, with the remainder being urbanized/suburbanized land [Li *et al.* 2006]. Residence times of water within the reservoir average 50-60 days [Li *et al.* 2006] but are highly variable and can be much shorter during high-influx events (daily-averaged minimum of 0.87 days, maximum of 13,877 days during 2013) [Clercin and Druschel 2019]. Maximum depths reach 11 to 13 m in the southern basin nearest the dam but average 3 to 5 m in most other portions. ECR is a dimictic water body with seasonal stratification of the water column; mixings usually occur in April or May and October each year, while the water column becomes thermally stratified between June and September.

The reservoir, like many freshwater bodies around the globe, is suffering from eutrophication driven by a large excess of nutrients within the watershed [Li *et al.* 2006]. Because of this, ECR has been experiencing annual algal blooms for well over a decade, with the first recorded measurement of T&O compounds dating back to at least 2000 [Pascual and Tedesco 2004]. During 2018 Indiana's Department of Environmental Management (IDEM) classified ECR as being mesotrophic with an unknown change in status (worsening or improving) through time [IDEM 2018].

Past studies of ECR

ECR has been the subject of several studies, a handful of which have been published, and investigations of T&O events at the reservoir are limited. One early study by Pascual and Tedesco (2004) examined physical and chemical conditions in the reservoir across four sampling sites as they relate to T&O concentrations and

phytoplankton abundance. Samples were collected before and after two copper algaecide treatments during 2003 to assess the efficacy and impact on phytoplankton growth. Physical conditions in the reservoir were found to vary seasonally with portions of ECR becoming thermally stratified during summer and fall. Changes in the system were accompanied by shifts in nutrient availability as well, with evidence of bloom growth switching between C-limited, N-limited, and P-limited states throughout the season. The two treatments varied greatly in effectiveness seemingly based on bloom speciation and nutrient conditions. The first algaecide application occurred when phytoplankton abundance was minimal and the population composed of diatoms and green algae. The treatment had little impact, likely tied to phytoplankton community composition and unfavorable nutrient conditions. The second treatment occurred when the community included a substantial portion of cyanobacteria and water column nutrient conditions were primed for exponential bloom growth. This was far more successful and effectively halted bloom progression, cutting algal biomass by several fold and leading to a decrease in T&O levels. The treatment appeared to trigger a shift from a P-abundant to P-limited nutrient environment which promoted a more diverse phytoplankton community structure less prone to nuisance T&O compounds.

As early as the Pascual and Tedesco (2004) study, GSM and MIB had been identified as the nuisance T&O compounds affecting ECR seasonally. Clercin and Druschel (2019) sampled across multiple depths from a single site located in the deepest section of ECR during 2013, with a focus on bacterial genetics and a detailed examination of bloom microorganisms' identities. Statistical analyses revealed that Actinomycetes bacteria were tied to MIB concentrations while several cyanobacterial

genera were linked to GSM. An algaecide treatment during the sampling period caused an immediate cessation of elevated GSM and severely impacted the cyanobacterial community. MIB levels were slower to respond and actinobacteria like *Streptomyces* were virtually unaffected. Conditions in the reservoir changed from being well mixed to thermally-stratified following this treatment. T&O production appeared to continue during this time but was confined to the bottom of the water column, and many microorganisms saw a slow resurgence in their population. Late season mixing reintroduced GSM and MIB into the upper water column as cyanobacteria grew rapidly, but T&O concentrations never approached the early season peak in May. Evidence showed that, overall, T&O producers at ECR prefer cold, well-mixed, highly turbid, and nutrient rich conditions seen earlier in the year, and that low TN:TP nutrient abundances likely triggered heightened T&O levels.

A study containing greater spatial and temporal detail in its sampling is needed to confirm these prior results and identify any relationships between T&O concentrations and observable environmental parameters that may have eluded previous studies. Many groups studying phytoplankton and cyanobacteria have seen success scouring datasets for meaningful relationships using simple exploratory techniques such as Pearson and Spearman correlations or more powerful explanatory techniques like Principal Component Analysis (PCA) and Non-Metric Multi-Dimensional Scaling (NMDS) across a number of freshwater lakes and reservoirs [Cunha and Calijuri 2011, Xu *et al.* 2011, Fidlerová and Hlúbiková 2016]. Cunha and Calijuri (2011) used the Spearman Rank Correlation Test and found that TP explained a greater percentage of bloom biomass variance than TN for the subtropical reservoirs they were studying. Additionally, PCA

illustrated that allochthonous organic matter in the riverine portion of the river had a strong effect on water column conditions. Fidlerová and Hlúbiková (2016) used Pearson's correlations to find pairwise interactions and determine any multicollinearities before the data was fed into PCA. The calculated principal component axes were able to separate different Slovakian reservoirs from one another based on the degree of water degradation and overall volume or flow, showing important environmental differences between locations separated by only a few hundred kilometers. Xu *et al.* (2011) successfully used NMDS to show that samples from the various stages of a diatom bloom at the Three-Gorges Reservoir were linked to changes in environmental conditions and were separable from each other when grouped by these stages. Even more complex analyses utilizing machine learning models, such as Boosted Regression Trees (BRT), have been made widely available using modern computers and open-source software while providing well-documented qualitative and predictive capabilities unmatched by simpler mathematic techniques [Elith *et al.* 2008].

The study presented here is an application of the aforementioned analyses to a large and complex dataset generated from samples collected at ECR. The focus of these endeavors is to interrogate and pick apart relationships between GSM or MIB and varying forms of environmental measurements applicable to shallow freshwater lakes and reservoirs worldwide. Such knowledge will be instrumental in forecasting future T&O events and informing best management practices for the reservoir. Our study includes a denser sampling regimen, exchanging depth profile data for a more temporally and spatially expansive dataset of the near-surface waters with emphasis on environmental conditions. By sampling in this way, important spatio-temporal information is captured

and included in the data, providing insights into the spatial and temporal dynamics of T&O generation not afforded by past studies.

Methodology:

Sampling regimen

Several hundred water samplings were performed during 2018 from the end of April through the middle of October, weekly, as weather permitted, across 14 locations in ECR (Figure 1.1). These locations were mostly focused on the northern basin of the reservoir where the watershed drains, with 12 locations spread across the northern area. The final two locations were next to the water treatment intake (at the northern most point of the southern basin, next to the outflow of the bottleneck separating both basins) and in the southern, deeper, section of the reservoir, near the ECR dam. Samples were gathered off a small boat between 10AM and 1PM and immediately sub-sectioned for analysis. Each day of sampling, water samples were collected using a Van Dorn sampler at 1m depth. Sub-samples included an unfiltered fraction (TP, Fe, Mn analyses) in a 50mL Falcon tube, another fraction (TDP, SRP, Ca, Mg, K, Na, DN analyses) filtered using a .22 µm GVWP filter in a 50mL Falcon tube, a third fraction of 50 mL spiked with ~1 mL of Lugol's Solution and stored in an amber glass vial for algae cell counts, another unfiltered fraction placed in an opaque 1L plastic bottle (Chl-a, PC, TSS analyses,) an unfiltered fraction placed in a 250mL amber bottle with septum for T&O compound analysis, and finally a sample for genetic analysis collected by filtering 50 mL of reservoir water through a Sterivex filter. Sub-samples were immediately placed on ice in a cooler until they could be transported to the lab and refrigerated.

Hyperspectral remote sensing reflectance below the water surface (rrs) spectra were measured using a dual head Ocean Optics (OO) USB4000 system composed of two spectroradiometers (351-1047 nm, 0.2 nm spectral resolution, 3645 bands). One spectroradiometer was mounted on the end of a 2 m long pole pointed directly upwards, measuring the downwelling irradiance (Ed), while the other spectroradiometer equipped with a 25˚ field-of-view optical fiber was dipped approximately 5 cm below the water surface to measure below-surface upwelling radiance (Lu) at nadir. The final rrs were calculated following the procedure described by Gitelson *et al.* (2007).

Chemical and physical analyses

Metals analysis was performed using an inductively coupled plasma – optical emission spectrometry (ICP-OES) instrument [Perkin Elmer Optima 7000 DV OES] focused on Fe, Mn, Ca, Mg, K, and Na. Fe and Mn were analyzed using unfiltered samples, while Ca, Mg, K, and Na were analyzed using filtered samples following the EPA method 200.7. Both unfiltered and filtered samples were spiked with HNO_3 to a concentration of approximately 2% and mixed well before experimentation. The analysis was carried out pursuant to the EPA method 200.7 for metals analysis in water.

Soluble reactive phosphorus (SRP) was determined by analyzing filtered samples, using the D'Angelo *et al.* (2001) method with the Malachite Green addendum modified to work with microplate-sized sample volumes. Samples were run in quadruplicate using a 96-well microplate and BioTek Epoch 2 microplate reader [D'Angelo *et al.* 2001].

Total phosphorus (TP) and total dissolved phosphorus (TDP) were measured using the Hach Total Phosphate Reagent Set (#2742645) following method 8910. TP

analysis was performed on unfiltered samples and TDP on their filtered counterparts, with absorbance measurements being made using the BioTek Epoch 2 microplate reader.

Dissolved nitrate (DN) was analyzed using filtered sub-samples and an ion chromatography system [Thermo Scientific Dionex ICS-1100 with AS-DV Autosampler, Dionex IonPac AS14 4x250 mm IC column and ASRS 300 4mm Suppressor, using 3.5mM Na_2CO_3/1mM $NaHCO_3$ eluent,] with samples ran in triplicate. Analysis was conducted following EPA method 300.00.

Algae cell counts were conducted using microscopy with a Nagoette Bright Line Hemacytometer stage (Hausser Scientific, 0.500 mm stage depth) and an Olympus BX53 microscope (100x, 200x, and 400x magnifications) on unfiltered water samples treated with 1-2% Lugol's solution. Microorganisms were sorted and characterized by phylum rather than species to facilitate sample throughput. The process followed the methods outlined by IUPUI's Center for Earth and Environmental Science (CEES) in their protocols for water monitoring [Clercin 2010].

Total Suspended Solids was measured using a simple gravimetric analysis, following EPA method 160.2. A GF/F filter was dry-ashed (103-105˚C) and weighed prior to filtering 300 mL of sample. The filter was then re-dried and weighed once more to determine the concentration of suspended solids in the original solution.

Chlorophyll A (Chl-*a*) content was measured following the SOP established by Arar (1997) and using a Shimadzu UV-2401PC UV-Vis Spectrophotometer. In short, 100 mL of sample was filtered using a GF/F filter and the filter placed in a Falcon tube containing 10 mL of 90% acetone then allowed to steep in a refrigerator overnight. The

solution was collected and centrifuged, then the supernatant placed in a quartz cuvette for analysis [Arar 1997].

The phycocyanin (PC) analysis used is an amalgamation of several protocols [Siegelman *et al.* 1978, Sarada *et al.* 1999, Horváth *et al.* 2013] adapted for ECR samples. 100 mL of sample was passed through a GF/F filter and the filter placed in a Falcon tube with 10 mL of a 0.1M phosphate buffer solution (10.5 g/L NaH_2PO_4, 4.4 g/L $Na_2HPO_4*12H_2O$ at pH 6). The sample underwent three freeze-thaw cycles (frozen for 2 hours, thawed, incubated in dark refrigeration for 24 h) before the solution was centrifuged. The supernatant was then collected and analyzed using a Turner Designs TD-700 Fluorometer to determine the concentration of PC.

The T&O compounds GSM and MIB were analyzed courtesy of the White River Water Treatment Plant labs of Citizens Energy Group. Samples were collected in 250 mL amber glass bottles with septum caps and delivered to the lab within an hour of field sampling concluding. Concentrations of GSM and MIB were quantified using Head-Space Solid-Phase Micro-Extraction (HS-SPME) coupled with a Gas Chromatography-Mass Spectrometry (GC-MS) instrument following the method SM 6040D [APHA 2000].

Genetics data was measured in IUPUI's biology department using 16S qPCR assays as well as qPCR assays for MIB and GSM synthases on gathered water samples. Genetic material was attained from Sterivex filters collected at each sampling site and frozen for storage within several hours. The analysis followed the procedure detailed in Koltsidou 2019, based partly on metagenomics data from ECR (Clercin and Druschel 2019), utilizing primers developed for detecting MIB/GSM synthases of cyano- and

actinobacterial organisms [Koltsidou 2019]. In short, DNA in the water samples was extracted using the DNeasy PowerWater Sterivex Kit from Qiagen. Lysis buffer was added to each Sterivex unit and mixed, then the lysate transferred to a bead beating tube for further lysis. Following this process DNA was captured on an MB Spin Column, washed, and the purified DNA eluted using a sterile elution buffer. This solution was immediately stored at -20˚C for later qPCR analyses. 16S assays as well as *geoA* and *MIBS* assays were conducted on aliquots of the purified DNA using a mix of TaqMan Universal PCR Master Mix, bovine serum albumin, and ECR-specific primers and probes designed for each experiment.

Data compilation

In addition to experimental and hyperspectral data, information about weather and hydrologic conditions was pooled together from online data repositories. Historical weather information, collected by the Eagle Creek Airpark KEYE station and the National Weather Service, was compiled and taken from *wunderground.com* for the months spanning the 2018 sampling season (data sourced from and consistent with National Weather Service station at the airpark). Hydrologic data was compiled using the USGS National Water Information System from gages 03353200 ("super"-gage located within the watershed) and 03353451 (located at the outflow of the dam). A substantial combined dataset (n = 333, variables = 132) containing information for every 2018 ECR sample was constructed.

Statistical analyses

The first analyses used were Pearson's (r) and Spearman's (r_s) correlation coefficients, which are pairwise tests intended to identify the direction, strength, and

significance of the relationship between any two variables. Pearson's correlation evaluates linear relationships and assumes that the variables are continuous, normally distributed, and do not contain any outliers, the latter two assumptions likely being violated by several variables in the ECR dataset. Spearman's correlation is a non-parametric test that only assumes observations are independent of each other and that the variables have a monotonic relationship, making it more appropriate for the given data. However, non-parametric tests inherently use less information (Spearman's is based on rank values whereas Pearson's can use raw data) and are considered less powerful.

PCA and NMDS analysis were also carried out on the 2018 data, both of which are multivariate tests often used to examine which variables are driving changes in the overall dataset. PCA is known as an R-mode analysis, that is, it is variable-focused and reliant on the interactions between variables. In broad terms the samples are plotted in n-dimensional space, with each variable contained within its own axis, and matrix algebra employed to generate a new coordinate system where each axis is a principal component. These principal components are effectively new variables composed of some combination of the original variables so that each component, beginning with the first, explains the most variance in the data as possible. The goal is to reduce the data to a more interpretable number of dimensions, preferably two or three, while preserving the differences between the samples as best as possible. PCA assumes that there are linear relationships between all of the variables and that these relationships are adequately strong, and it is sensitive to variables with a large number of zeros or significant outliers. In general, if the first two to three dimensions are able to capture 70% or more of the

variance then the results are considered strong and informative [McKillup and Dyar 2010].

NMDS is a Q-mode analysis which focuses on similarities between samples. It is related to PCA in much the same way Spearman's correlation is to Pearson's, and is reliant on monotonic relationships between variables. The analysis works by calculating a dissimilarity matrix between each possible pair of samples, then plotting all samples in a defined number of dimensions and recalculating dissimilarity. The original matrix is compared to the new one using a measure termed stress, and then all sample points adjusted to slightly better matching positions. This is done iteratively until a global minima in stress is found, resulting in a final plot of all samples. Stress values greater than 0.2 are considered very poor and likely to lead to false interpretations, while stress less than 0.1 is considered a good result and stress approaching 0.05 or less corresponds to an excellent result [Clarke 1993].

Finally, the data was fed into the "dismo" and "gbm" [Elith *et al.* 2008] R packages to generate a Boosted Regression Tree (BRT) model with MIB concentration as the response variable. A BRT model is essentially the combination of decision tree and gradient boosting techniques. The algorithm constructs an aggregate model by combining a large number of smaller, simpler, decision tree models in a stagewise fashion using random subsets of the data. Choices at each step are made with the intention of reducing the residual deviance in the response variable and the process is carried forward until the best fit possible is achieved. BRT models include no assumptions about variable distribution and are non-parametric, making them robust to missing values, outliers, and skewed data. The final model can make predictions using other datasets, or it can be

picked apart using a number of tools to gain qualitative information about variable relationships. The package can graph the fitted functions for each predictor and rank them by significance, allowing the user to see how the response variable is changing in relation to every predictor and how important said predictor is to the model. Similarly, the strength of pairwise interactions between predictors in the model can ranked to elucidate relationships not directly pertaining to the response.

Data imputation

A major problem with large datasets collected from the environment is missing or incomplete data due to any number of reasons. Analyses like pairwise correlations can easily overcome modest amounts of missing information by ignoring offending entries during each pairwise test, leaving subsequent tests unaffected. PCA and NMDS necessitate that every sample has information for every variable, which can quickly result in some variables and many samples being removed from the dataset. In the case of the ECR data entire sampling days are lost, shrinking the number of observations considerably ($n < 200$). Missing data can be imputed through simple methods like using the mean or median of a variable to fill in missing observations, but this can heavily skew the data and distort the results of PCA or NMDS.

More sophisticated imputation methods exist, each with their own strengths and weaknesses. For this study the "missMDA" package for R [Josse and Husson 2016] was used for both PCA and NMDS analysis as well as data imputation. A comparative study found that the imputation algorithm used by missMDA, termed iterative PCA (IPCA), was far superior to mean/median imputation and surpassed many other advanced methods across a number of scenarios [Dray and Josse 2015]. Furthermore, the ECR dataset was

analyzed both with and without imputation using PCA and NMDS and the results were very similar, but the imputed dataset gives more precision and allows the model to be extended over greater time and space (see Figures A1.1 and A1.2 for a comparison between non-imputed and imputed results).

References

Albarracín, V.H., A.L. Ávila, M.J. Amoroso, C.M. Abate. 2008. Copper removal ability by Streptomyces strains with dissimilar growth patterns and endowed with cupric reductase activity. FEMS Microbiol. Lett. 288: 141-148, doi: 10.1111/j.1574-6968.2008.01335.x

Andersen, I.M., T.J. Williamson, M.J. González, M.J. Vanni. 2020. Nitrate, ammonium, and phosphorus drive seasonal nutrient limitation of chlorophytes, cyanobacteria, and diatoms in a hyper-eutrophic reservoir. Limnol. Oceanogr. 65: 962-978, doi: 10.1002/lno.11363

Anderson, D. 2014. HABs in a changing world: a perspective on harmful algal blooms, their impacts, and research and management in a dynamic era of climactic and environmental change. Harmful Algae. 3-17.

Anderson, D.M., P.M. Gilbert, and J.M. Burkholder. 2002. Harmful Algal Blooms and Eutrophication: Nutrient Sources, Composition, and Consequences. Estuaries. 25: 704-726, doi: 140.182.176.10

APHA. 2000. Supplement to Standard Methods for the Examination of Water and Wastewater, 20th ed. APHA, AWWA, WPCF.

Arar, E.J. 1997. Method 446.0: In Vitro Determination of Chlorophylls a, b, c + c, and Pheopigments in Marine and Freshwater Algae by Visible Spectrophotometry. U.S. EPA. EPA/600/R-15/005.

Barker, T., H. MD. Irfanullah, and B. Moss. 2010. Micro-scale structure in the chemistry and biology of a shallow lake. Freshwater Biology. 55: 1145-1163, doi: 10.1111/j.1365-2427.2009.02339.x

Bentley, R., and R. Meganathan. 1981. GSM and Methylisoborneol Biosynthesis in Streptomycetes: Evidence for an Isoprenoid Pathway and its Absence in Non-differentiating Isolates. FEBS letters. 125: 220-22, doi: 10.1016/0014-5793(81)80723-5

Bishop, W.M., B.M. Johnson, J.H. Rodgers, Jr. 2014. Comparative responses of target and nontarget species to exposures of a copper-based algaecide. J. Aquat. Plant Manage. 52: 65-70

Carmichael, W. W. 2001. Health Effects of Toxin-producing Cyanobacteria: 'The Cyano-HABs'. Human and Eco. Risk Assess. 7: 1393-1407, doi: 10.1080/20018091095087

Carpenter, S.R., N.F. Caraco, D.L. Correll, R.W. Howarth, A.N. Sharpley, V.H. Smith. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications. 8: 559-568

Christophoridis, C., K. Fytianos. 2006. Conditions Affecting the Release of Phoshorus from Surface Lake Sediments. J. Environ. Qual. 35: 1181-1192, doi: 10.2134/jeq2005.0213

Clarke, K.R. 1993. Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology. 18: 117-143.

Clercin, N.A. 2010. Monitoring Standard for Freshwater Algae: Analysis and Enumeration of Planktic Microalgae. CEES Protocol Standards.

Clercin, Nicolas André. 2019. Origin and Fate of Odorous Metabolites, 2-Methylisoborneol and Geosmin, in a Eutrophic Reservoir. Department of Earth Sciences, Indiana University

Clercin, N.A. and G.K. Druschel. 2019. Influence of Environmental Factors on Off-Flavor Metabolite Production by Bacteria in a Eutrophic Reservoir. Water Resources Research. 55: 5413-30, doi: 10.1029/2018WR023651

Cunha, D.G.F. and M.d.C. Calijuri. 2010. Limiting factors for phytoplankton growth in subtropical reservoirs: the effect of light and nutrient availability in different longitudinal compartments. Lake and Reservoir Management. 27: 162-172, doi: 10.1080/07438141.2011.574974

D'Angelo, E., J. Crutchfield, M. Vandiviere. 2001. Rapid, Sensitive, Microscale Determination of Phosphate in Water and Soil. J. Environ. Qual. 30: 2206-09, doi: 10.2134/jeq2001.2206

Davidson, K. and others. 2012. Harmful algal blooms: How strong is the evidence that nutrient ratios and forms influence their occurrence? Estuarine, Coastal, and Shelf Science. 115: 399-413, doi: 10.1016/j.ecss.2012.09.019

Delwiche, C.C. 1970. The Nitrogen Cycle. Scientific American. 223: 136-147

Diaz, R.J. 2001. Overview of Hypoxia around the World. Journal of Environmental Quality. 30: 275-281.

Dodds, W. K., W. W. Bouska, J. L. Eitzmann, T. J. Pilger, K. L. Pitts, A. J. Riley, J. T. Schloesser, and D. J. Thornbrugh. 2008. Eutrophication of US Freshwaters: Analysis of Potential Economic Damages. Environ. Sci. Technol. 43: 12-19. doi: 10.1021/es801217q

Dray, S. and J. Josse. 2015. Principal Component Analysis with Missing Values: a Comparative Survey of Methods. Plan Ecol. 216: 657–667, doi: 10.1007/s11258-014-0406-z

Elith, J., J.R. Leathwick, T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology. 77: 802-13, doi: 10.1111/j.1365-2656.2008.01390.x

Fay, P. 1992. Oxygen Relations of Nitrogen Fixation in Cyanobacteria. Microbiological Reviews. 56: 340-73.

Filderová, D., and D. Hlúbiková. 2016. Relationships between benthic diatom assemblages' structure and selected environmental parameters in Slovak water reservoirs (Slovakia, Europe). Knowl. Manag. Aquat. Ecosyst. 417: 27, doi: 10.1051/kmae/2016014

Flynn, K.J. 2010. Do external resource ratios matter?: Implications for modelling eutrophication events and controlling harmful algal blooms. Journal of Marine Systems. 83: 170-180.

Gerber, N., and H. Lechevalier. 1965. GSM, an Earthy-smelling Substance Isolated from Actinomycetes. Applied Microbio. 13: 935-38.

Gitelson, A.A., J.F. Schalles, C.M. Hladik. 2007. Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. Remote Sensing of Environment. 109: 464-72, doi: 10.1016/j.rse.2007.01.016

Glibert, P.M., T.M. Kana, K. Brown. 2013. From limitation to excess: the consequences of substrate excess and stoichiometry for phytoplankton physiology, trophodynamics and biogeochemistry, and the implications for modeling. Journal of Marine Systems. 125: 14-28

Graham, J.L., G.M. Foster, T.J. Williams, M.D. Mahoney, M.R. May, K.A. Loftin. 2018. Water-quality conditions with an emphasis on cyanobacteria and associated toxins and taste-and-odor compounds in the Kansas River, Kansas, July 2012 through September 2016. U.S. Geological Survey Scientific Investigations Report. doi: 10.3133/sir20185089

Graham, J.L., K.A. Loftin, M.T. Meyer, A.C. Ziegler. 2010. Cyanotoxin Mixtures and Taste-and-Odor Compounds in Cyanobacterial Blooms from the Midwestern United States. Environ. Sci. Technol. 44: 7361-7368, doi: 10.1021/es1008938

Hamill, K. D. 2001. Toxicity in Benthic Freshwater Cyanobacteria (Blue-green Algae): First Observations in New Zealand. New Zealand Journal of Marine and Freshwater Research. 35: 1057-59, doi: 10.1080/00288330.2001.9517062

Harris, T. D., V. H. Smith, J. L. Graham, D. B. Van de Waal, L. P. Tedesco, and N. Clercin. 2016. Combined Effects of Nitrogen to Phosphorus Ratios and Nitrogen Speciation on Cyanobacterial Metabolite Concentrations in Eutrophic Midwestern USA Reservoirs. Inland Waters. 6: 199-210, doi: 10.5268/IW-6.2.938

Heisler, J. and others. 2008. Eutrophication and Harmful Algal Blooms: A Scientific Consensus. Harmful Algae. 8: 3-13, doi: 10.1016/j.hal.2008.08.006

Horváth, H., A.W. Kovács, C. Riddick, M. Présing. 2013. Extraction Methods for Phycocyanin Determination in Freshwater Filamentous Cyanobacteria and Their Application in a Shallow Lake. Eur. J. Phycol. 48: 278-86, doi: 10.1080/09670262.2013.821525

Indiana Department of Environmental Management. 2018. Indiana Integrated Water Quality Monitoring and Assessment Report.

Jenne, E.A. 1968. Controls on Mn, Fe, Co, Ni, Cu, and Zn concentrations in soils and water: the significant role of hydrous Mn and Fe oxides. Advances in Chemistry. 73: 337-387

Josse, J., and F. Husson. 2016. missMDA: A Package for Handling Missing Values in Multivariate Data Analysis. Journal of Statistical Software. 70.

Juttner, F. 1990. Monoterpenes and Microbial Metabolites in the Soil. Environ. Pollut. 68: 377-82.

Juttner, F., and S. B. Watson. 2007. Biochemical and Ecological Control of GSM and 2-methylisoborneol in Source Waters. Appl. Environ. Microbio. 73: 4395-406, doi: 10.1128/AEM.02250-06

Kansole, M. M.R. and T.-F. Lin. 2017. Impacts of Hydrogen Peroxide and Copper Sulfate on the Control of *Microcystis aeruginosa* and MC-LR and the Inhibition of MC-LR Degrading Bacterium *Bacillus* sp. Water. 9: 255, doi: 10.3390/w9040255

Koltsidou, Ioanna. 2019. Detection and Quantification of Taste and Odor Producing Bacteria in Eagle Creek Reservoir. Department of Biological Sciences, Purdue University.

Komatsu, M., M. Tsuda, S. Ōmura, H. Oikawa, and H. Ikeda. 2008. Identification and Functional Analysis of Genes Controlling Biosynthesis of 2-methylisoborneol. Proceedings of the National Academy of Sciences. 105: 7422-27, doi: 10.1073/pnas.0802312105

Krishnani, K. K., P. Ravichandran, and S. Ayyappan. 2008. Microbially Derived Off-flavor from GSM and 2-methylisoborneol: Sources and Remediation. Rev. Environ. Contam. Toxicol. 194: 1-27, doi: 10.1007/978-0-387-74816-0_1

Lanciotti, E., C. Santini, E. Lupi, and D. Burrini. 2003. Actinomycetes, Cyanobacteria and Algae Causing Tastes and Odours in Water of the River Arno Used for the Water Supply of Florence. Journal of Water Supply: Research and Technology – Aqua. 52: 489-500, doi: 10.2166/aqua.2003.0044

Lean, D.R.S. and others. 1987. The Effects of Changes in Both the Abundance of Nitrogen and Phosphorus and Their Ratio on Lake Okaro Phytoplankton, with Comment on Six Other Central Volcanic Plateau Lakes. Marine and Freshwater Research. 21: 539-42, doi: 10.1080/00288330.1987.9516257

Leschine, S., K. Holwell, and E. Canale-Parola. 1988. Nitrogen Fixation by Anaerobic Cellulolytic Bacteria. Science. 242: 1157-59, doi: 10.1126/science.242.4882.1157

Li, L., D.L. Pascual, L.P. Tedesco, K.L. Randolph, R.E. Sengpiel, B.E. Hall, J.S. Wilson. 2006. Developing a Survey Tool for the Rapid Assessment of Blue-Green Algae in Central Indiana’s Reservoirs. IDNR LARE Report, CEES Publication 2006-1.

Li, L., N. Wan, N. Gan, B. Xia, and L. Song. 2007. Annual Dynamics and Origins of the Odorous Compounds in the Pilot Experimental Area of Lake Dianchi, China. Water Sci. and Technol. 55: 43-50, doi: 10.2166/wst.2007.160

Lijklema, L. 1980. Interaction of Orthophosphate with Iron(III) and Aluminum Hydroxides. American Chemical Society. 14: 537-541

Łukawska-Matuszewska, K., R.D. Vogt, R. Xie. 2013. Phosphorus pools and internal loading in a eutrophic lake with gradients in sediment geochemistry created by land use in the watershed. Hydrobiologia. 713: 183-197

Ma, Z. M., Y. Niu, P. Xie, J. Chen, M. Tao, and X. W. Deng. 2013. Off-flavor Compounds from Decaying Cyanobacterial Blooms of Lake Taihu. J. Environ. Sci-China. 25: 495-501, doi: 10.1016/S1001-0742(12)60101-6

McKillup, S. and M.D. Dyar. 2010. Geostatistics Explained: an Introductory Guide for Earth Scientists, Cambridge University Press.

Miguéns, D., and E. Valério. 2015. The Impact of Some Microcystins on the Growth of Heterotrophic Bacteria from Portuguese Freshwater Reservoir. Limnetica. 34: 215-226, doi: 10.23818/limn.34.17

Nyhus, K.J. and E.S. Jacobson. 1999. Genetic and Physiologic Characterization of Ferric/Cupric Reductase Constitutive Mutants of Cryptococcus neoformans. Infection and Immunity. 67: 2357-2365.

Ömür-Özbek, P., J.C. Little, A.M. Dietrich. 2007. Ability of humans to smell GSM, 2-MIB, and nonadienal in indoor air when using contaminated drinking water. Water Sci. and Tech. 55: 249-56, doi: 10.2166/wst.2007.186

Paerl, H. W., and J. Huisman. 2009. Climate Change: A Catalyst for Global Expansion of Harmful Cyanobacterial Blooms. Environ. Microbio. Reports. 1: 27-37, doi: 10.1111/j.1758-2229.2008.00004.x

Paerl, H. W., R. S. Fulton, P. H. Moisander, and J. Dyble. 2001. Harmful Freshwater Algal Blooms, with an Emphasis on Cyanobacteria. The Scientific World Journal. 1: 76-113, doi: 10.1100/tsw.2001.16

Pascual, D.L and L.P. Tedesco. 2004. Eagle Creek Reservoir: Responses to Algaecide Treatment. CEES Publication 2004-02.

Perkins, R.G., and others. 2019. Managing taste and odour metabolite production in drinking water reservoirs: The importance of ammonium as a key nutrient trigger. Journal of Environmental Management. 244: 276-284, doi: 10.1016/j.jenvman.2019.04.123

Petterson, K. 1998. Mechanisms for internal loading of phosphorus in lakes. Hydrobiologia. 373/374: 21-25.

Pierzynski, G.M., R.W. McDowell, J.T. Sims. 2005. Chemistry, Cycling, and Potential Movement of Inorganic Phosphorus in Soils. doi: 10.2134/agronmonogr46.c3

Pinel-Alloul, B. 1995. Spatial heterogeneity as a multiscale characteristic of zooplankton community. Hydrobiologia. 300/301: 17-42.

Pinel-Alloul, B., and A. Ghadouani. 2007. The Spatial Distribution of Microbes in the Environment, Springer.

Podduturi, R., N. Clausen, N.B. Petersen, N. O.G. Jørgensen, M.A. Petersen. 2018. Continuous collection of volatiles produced by *Streptomyces* grown on oatmeal agar by headspace extraction and GC-MS. doi: 10.3217/978-3-85125-593-5-100

Ramsing, N. B., H. Fossing, T. G. Ferdelman, F. Andersen, and B. Thamdrup. 1996. Distribution of Bacterial Populations in a Stratified Fjord (Mariager Fjord, Denmark) Quantified by In Situ Hybridization and Related to Chemical Gradients in the Water Column. Applied and Environ. Microbio. 62: 1391-1404.

Sarada, R., M.G. Pillai, G.A. Ravishankar. 1999. Phycocyanin from Spirulina sp.: Influence of Processing of Biomass on Phycocyanin Yield, C-Phycocyanin Extraction from Spirulina platensis Wet Biomass Analysis of Efficacy of Extraction Methods and Stability Studies on Phycocyanin. Process Biochemistry. 34: 795.

Shatwell, T., J. Koehler, and A. Nicklisch. 2008. Warming Promotes Cold-adapted Phytoplankton in Temperate Lakes and Opens a Loophole for Oscillatoriales in Spring. Global Change Biology. 14: 2194-2200, doi: 10.1111/j.1365-2486.2008.01630.x

Shen, X., H. Zhang, X. He, H. Shi, C. Stephan, H. Jiang, C. Wan, and T. Eicholz. 2019. Evaluating the treatment effectiveness of copper-based algaecides on toxic algae Microsystis aeruginosa using single cell-inductively coupled plasma-mass spectrometry. Analytical and Bioanalytical Chemistry. 411: 5531-5543.

Siegelman, H.W., J.H. Kycia, J.A. Hellebust, J.S. Craigie. 1978. Algal Biliproteins. Handbook of Phycological Methods: Physiological and Biochemical Methods. 72-78.

Šimek, K., M. Comerma, J.-C. García, J. Nedoma, R. Marcé, and J. Armengol. 2011. The Effect of River Water Circulation on the Distribution and Functioning of Reservoir Microbial Communities as Determined by a Relative Distance Approach. Ecosystems. 14: 1-14, doi: 10.1007/s10021-010-9388-4

Smayda, T.J. 1997. Harmful phytoplankton blooms: their ecophysiology and general relevance. Limnol. Oceanogr. 42: 1137-1153.

Smil, V. 2000. Phosphorus in the Environment: Natural Flows and Human Interferences. Annu. Rev. Energy Environ. 25: 53-88.

Smith, K.S. 1999. Metal sorption on mineral surfaces- An overview with examples relating to mineral deposits; in Plumlee, G.S., and Logsdon, M.J. The environmental Geochemistry of Mineral Deposits, Part A. Processes, Techniques, and Health Issues: Society of Economic Geologists, Reviews in Economic Geology. 6A: 161-182.

Smith, L., M.C. Watzin, G. Druschel. 2011. Relating sediment phosphorus mobility to seasonal and diel redox fluctuations at the sediment-water interface in a eutrophic freshwater lake. Limnol. Oceanogr. 56: 2251-2264, doi: 10.4319/lo.2011.56.6.2251

Smith, V.H., G.D. Tilman, J.C. Nekola. 1999. Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. Environmental Pollution. 100: 179-196.

Søndergaard, M., B. Riike, B. Bullet, E. Jeppesen. 2013. Persistent internal phosphorus loading during summer in shallow eutrophic lakes. Hydrobiologia. 710: 95-107.

Søndergaard, M., J.P. Jensen, E. Jeppesen. 2003. Role of sediment and internal loading of phosphorus in shallow lakes. Hydrobiologia. 506-509: 135-145.

Srinivasan, R., and G. A. Sorial. 2011. Treatment of Taste and Odor Causing Compounds 2-methyl Isoborneol and GSM in Drinking Water: A Critical Review. J. Environ. Sci-China. 23: 1-13, doi: 10.1016/S1001-0742(10)60367-1

Stewart, W.D.P., F.R.S. T. Preston, H.G. Peterson, N. Christofi. 1982. Nitrogen cycling in eutrophic freshwaters. Phil. Trans. R. Soc. Lond. 296: 491-509

Su, M. and others. 2015. MIB-producing Cyanobacteria (Planktothrix sp.) in a Drinking Water Reservoir: Distribution and Odor Producing Potential. Water Research. 68: 444-53, doi: 10.1016/j.watres.2014.09.038

Tedesco, L.P., D.L. Pascual, L.K. Shrake, R.E. Hall, L.R. Casey, P.G.F. Vidon, F.V. Hernly, K.A. Salazar, R.C. Barr. 2005. Eagle Creek Watershed Management Plan: An Integrated Approach to Improved Water Quality. Eagle Creek Watershed Alliance, CEES Publication 2005-07, IUPUI, Indianapolis, 182p.

Tyc, O., and others. 2017. Exploring bacterial interspecific interactions for discovery of novel antimicrobial compounds. Microbial Biotechnology. 10: 910-925, doi: 10.1111/1751-7915.12735

Valiela, I. and others. 1992. Couplings of Watersheds and Coastal Waters: Sources and Consequences of Nutrient Enrichment in Waquoit Bay, Massachusetts. Estuaries. 15: 443-457

Vitousek, P.M and others. 1997. Human alteration of the global nitrogen cycle: sources and consequences. Ecological Applications. 7:737-750

Watson, S. B., P. Monis, P. Baker, and S. Giglio. 2016. Biochemistry and Genetics of Taste-and odor-producing Cyanobacteria. Harmful Algae. 54: 112-27, doi: 10.1016/j.hal.2015.11.008

Watson, S.B. 2003. Cyanobacterial and eukaryotic alagal odour compounds: signals or by-products? A review of their biological activity. Phycologia. 42: 332-350

Wells, M.L., and others. 2020. Future HAB science: Directions and challenges in a changing climate. Harmful Algae. 91, doi: 10.1016/j.hal.2019.101632

Wnorowski, A. 1992. Tastes and Odours in the Aquatic Environment: A Review. Water S. A. 18: 203-14.

Xu, Hai, Hans W. Paerl, Boqiang Qin, Guangwei Zhu, Guang Gao. 2010. Nitrogen and Phosphorus Inputs Control Phytoplankton Growth in Eutrophic Lake Taihu, China. Limnol. Oceanogr. 55: 420-32, doi: 10.4319/lo.2010.55.1.0420

Xu, Y., Q. Cai, L. Ye, M. Shao. 2011. Asynchrony of spring phytoplankton response to temperature driver within a spatial heterogeneity bay of Three-Gorges Reservoir, China. Limnologica. 41: 174-180, doi: 10.1016/j.limno.2010.10.004

Chapter 2 – Taste and Odor event dynamics of a Midwestern freshwater reservoir

Eagle Creek Reservoir, located in the Midwestern U.S., is a freshwater limnic system plagued by seasonal Harmful Algal Blooms (HABs) which generate water-fouling Geosmin (GSM) and 2-Methylisoborneol (MIB) Taste and Odor (T&O) compounds. Past investigations of T&O event dynamics have identified *Actinomycetes* as responsible for MIB production and several genera of cyanobacteria for GSM production [Clercin 2019]. During 2018 a temporally and spatially expansive sampling regimen of the reservoir was carried out and a battery of biological, chemical, physical, and hyperspectral measurements made. The resulting data was analyzed using time series, cross-correlation, lag time, and multivariate analyses as well as machine learning algorithms to interrogate potential relationships between HABs, T&O events, and environmental parameters. The results show that local weather and watershed conditions exert significant control over the state of the reservoir and the behavior of the algal community. GSM and MIB peaked during early May under well-mixed, cold, and nutrient-rich water column conditions, then declined under summer thermal stratification before making a small resurgence during late season mixing. Bloom die-off and decay was effectively ruled out as a mechanism controlling T&O concentrations, and no links were found between T&O concentrations and algal biomass. Strong evidence was found that GSM/MIB concentrations were a response by certain microbes to changing nutrient conditions within the reservoir, and it was determined that nutrient fluxes from the watershed 30-40 days prior to peak T&O concentrations are likely instrumental in the development of the slow-growing microbes that produce these compounds.

Introduction:

HAB prevalence and environmental impact

Harmful Algal Blooms (HABs) are a pronounced and growing problem for most freshwater bodies around the globe, the rise in the frequency and intensity of HABs is thought to be the result of degraded water quality from increase nutrient pollution, referred to as eutrophication [Paerl *et al.* 2001, Heisler *et al.* 2008], in combination with increasing temperatures driven by global warming/climate change [Shatwell *et al.* 2008, Paerl and Huisman 2009]. HABs have been found to cause significant ecological changes which may affect the overall ecosystem [Miguéns and Valério 2015]. In some localities HABs are hazardous to human health as they can be responsible for producing a wide range of neuro- and hepatotoxins [Carmichael 2001, Hamill 2001]. HABs are capable of depleting the available oxygen in the water column and creating anoxic conditions, which in turn can cause benthic mortalities and/or fish kills [Smith *et al.* 1999, Diaz 2001, Anderson *et al.* 2002]. HAB events can also be linked to the production of secondary metabolites, including a class of Volatile Organic Compounds (VOCs) referred to as Taste and Odor (T&O) compounds, considered to be non-toxic and largely unrelated to neurotoxin production in limnic systems [Graham *et al.* 2010]. These molecules are easily volatilized into the gaseous form, and human senses can detect them even in exceedingly low concentrations on the range of 10-20 ng/L [Ömür-Özbek *et al.* 2007]. This is a problem especially for drinking water as these T&O molecules have an easily observable negative impact on overall water quality [Srinivasan and Sorial 2011] which can cause significant economic damage. T&O events can be remediated through a battery of water treatment techniques (typically absorption to granular activated carbon (GAC) or

powdered activated carbon (PAC), advanced oxidative processes using UV/ozone/H_2O_2, or biologically activated filter mediums) [Srinivasan and Sorial 2011] or through algaecide application to the lake or reservoir [Pascual and Tedesco 2004], but with limited and varying efficacy often tied to the speciation of microorganisms within the blooms [Bishop *et al.* 2014, Kansole and Lin 2017, Shen *et al.* 2019]. For example, there have been repeated observations at Eagle Creek Reservoir (ECR) that cyanobacteria are far more susceptible to algaecide treatment than actinobacteria [Pascual and Tedesco 2004, Clercin and Druschel 2019]. Most algaecides, including those used at ECR, are copper-based and the differences in response to algaecide treatment between cyanobacteria and actinobacteria are likely due to each microorganism's ability to internally regulate copper concentrations. Copper algaecide primarily functions by elevating these copper concentrations to a cytotoxic level within the cell, where the element inhibits photosystem II activity and electron transport systems of cyanobacteria [Shen *et al.* 2019]. This leads to eventual cell lysis and death. Certain actinobacteria like *Streptomyces* have demonstrated a profound efficiency at eliminating copper from themselves and were found to possess cupric reductase enzymes [Albarracín *et al.* 2008]. Actinobacterial resistance to copper-based algaecides is probably linked to this ability to regulate copper transport across the cell membrane, while cyanobacterial susceptibility stems from their lack of such a mechanism.

T&O producers and T&O events

Of the T&O compounds, two specifically, Geosmin (GSM) and 2-Methylisoborneol (MIB), are the main culprits in surface freshwaters affected by HAB T&O events [Lanciotti *et al.* 2003, Ma *et al.* 2013]. GSM and MIB are produced by algal

species mostly within Actinobacteria [Gerber and Lechevalier 1965, Juttner 1990] and Cyanobacteria phyla [Juttner and Watson 2007, Li *et al.* 2007, Su *et al.* 2015]. Specific microorganisms and their genetic sequences responsible for T&O production have been identified [Bentley and Meganathan 1988, Komatsu *et al.* 2008, Watson *et al.* 2016] and the biosynthesis of both molecules is summarized by three main pathways. These are the 2-methylerythritol-4-phosphate (MEP), mevalonate (MVA), and leucine pathways, though MEP is the major biosynthetic route for many bacteria and leucine is thought to only play a minor role [Juttner and Watson 2007]. The MEP pathway has been observed in *Streptomyces* bacteria and the necessary genes found in several cyanobacteria species as well. GSM and MIB have specific synthases that determine the eventual product of these pathways, also present in species representing both Actinobacterial and Cyanobacterial groups [Kuzuyama 2002, Rodriguez-Concepción and Boronat 2002, Cane *et al.* 2006, Jiang *et al.* 2007, Giglio *et al.* 2008, Wang *et al.* 2008, Asquith *et al.* 2013]. The MVA pathway has been observed in *Streptomyces* too, with work by Bentley and Meganathan (1981) suggesting it was the predominant pathway for GSM production. Later evidence did not support MVA over MEP as the dominant mechanism, but showed MVA genetics were also present in some cyanobacteria. The purpose of T&O molecule production is not fully understood, but they are considered secondary metabolites and therefore unlikely to influence the growth or reproduction of their producers. It has been suggested that these chemicals influence microbial community dynamics and serve as either infochemicals used for communication or as antimicrobial agents used to fight competitors [Watson *et al.* 2003, Tyc *et al.* 2017].

The required circumstances for T&O compounds to be upregulated and expressed are not well understood. HABs and the T&O events they can generate are often a seasonal occurrence [Ramsing *et al.* 1996, Li *et al.* 2007, Su *et al.* 2015] and the two chemicals can be found both during HAB events and afterwards, being released with the decay of the bloom's microorganisms [Ma *et al.* 2013]. Past work pointed to overall bloom biomass and development controlling GSM concentrations rather than changing biosynthesis rates in response to environmental conditions [Wnorowski 1992]. However, lab experiments showed that GSM and MIB production rates were linked to biomass for certain *Streptomyces* species but not others [Podduturi *et al.* 2018]. Furthermore, a comprehensive multi-year study of the Kansas River found no relationships between GSM/MIB and algal biomass, cyanobacterial abundance, or actinomycetes abundance [Graham *et al.* 2018].

Role of nutrients and eutrophication

The growth and proliferation of HABs has been studied extensively and it is generally established that nitrogen (N) and phosphorus (P) are the limiting nutrients for phytoplankton growth in freshwater lakes [Lean *et al.* 1987, Xu *et al.* 2010]. The importance of each nutrient can vary from system to system, as several microorganisms commonly found in blooms have demonstrated N-fixation capabilities [Leschine *et al.* 1988, Fay 1992]. The molecular form of both nutrients and their respective sources is crucial to understanding the dynamics of nutrient-bloom interactions for an individual freshwater body. N and P can be derived from any number of natural or anthropogenic sources, leading to external and/or internal nutrient loading in a system [Xu *et al.* 2010]. Human activity within watersheds has greatly increased the rate/amount of nutrients

transported downstream, and in general increased urbanization is directly correlated with increased nutrient pollution [Valiela *et al.* 1992]. Point sources in watersheds include mine drainage/run-off and sewage outflows though these are often less important than non-point sources, which include fertilizer/manure application/run-off and fossil fuel combustion [Anderson *et al.* 2002]. The over-application of fertilizer and manure to fields, along with high density animal husbandry, tend to enrich soils in nutrients which can then be transported through leaching or soil erosion [Carpenter *et al.* 1998]. A firm grasp of these nutrient dynamics as they relate to a particular system is often key to knowing the likelihood and timing of HABs.

The impact human activity has had on the global N cycle is substantial and nitrogen's availability and mobility have been greatly enhanced through anthropogenic modification [Vitousek *et al.* 1997]. A majority of nitrogen is contained within a massive and virtually unavailable pool of nitrogen gas (N_2) in the Earth's atmosphere. A small portion of N_2 is converted through biologically mediated processes to more bioavailable chemical forms [Delwiche 1970]. This crucial step in the cycle, termed N fixation, converts N_2 to ammonia which can then undergo further conversion through nitrification to form nitrites and nitrates. Human activity has effectively doubled the speed of this transfer of N from the atmospheric pool to land and water, by industrial N fixation (fertilizer creation and application), fossil fuel combustion (production of nitrogen oxides), extensive sowing of N-fixing crops (leguminous and forages), and the burning/clearing of biomass [Vitousek *et al.* 1997]. This accelerated N fixation drives eutrophication of waters within a catchment or watershed; a British study using ^{15}N

isotope spiked fertilizer showed that over 90% of the added ^{15}N could be found within algal biomass during peak blooms [Stewart *et al.* 1982].

In many aquatic systems focus is often placed on P sources, dynamics, and cycling due to its relative scarcity when compared to N. P can be found in the land or water as part of a complex cycle which is generally less studied than the N cycle [Smil 2000]. External inputs of P into a body of water stem from a number of point and non-point sources [Anderson *et al.* 2002] like N, but lake sediments can serve as a large and dynamic reservoir for P leading to internal loading of the nutrient in some systems [Søndergaard *et al.* 2003]. Seasonal hypoxia and/or thermal stratification can lead to changes in pH and redox state within the system causing the reductive dissolution of iron oxides, this dissolution releases sorbed P and any P that had been incorporated into the mineral's structure, resulting in fluxes of P into the overlying water column [Pettersson 1998, Christophoros and Fytianos 2006, Smith *et al.* 2011]. The process of P introduction to a freshwater lake, its sedimentation/mineralization/immobilization in the system, and seasonal hypoxia and/or thermal stratification causing release can mobilize P in fluxes several times greater than externally loaded phosphorus concentrations [Søndergaard *et al.* 2013].

Variability in nitrate, ammonium, and phosphorus ratios was strongly correlated to seasonal algal community trends and shifts in the state of nutrient limitation [Andersen *et al.* 2020]. Across Midwestern reservoirs low TN:TP ratios were seen to favor cyanobacterial dominance and growth while low NO_3:NH_3 ratios were associated with increased T&O production [Harris *et al.* 2016]. Other studies have found similar associations between low NO_3:NH_4 ratios and heightened T&O production [Perkins *et al.*

2019]. The two general assumptions used in analyzing these ratios are that 1. low N:P (N-limited) or high N:P (P-limited) ratios are indicative of the state of nutrient limitation and two, or that 2. different nutrient forms have varying bioavailability and ratios such as NO_3:NH_3 affect community composition.

HAB distributions

The spatio-temporal distribution of HABs is not as well studied as nutrient importance in phytoplankton growth, and the role of nutrient concentrations versus other environmental factors is still highly debated [Flynn 2010, Davidson *et al.* 2012, Glibert *et al.* 2013, Anderson 2014]. Several studies have shown that local hydrology [Šimek *et al.* 2011] and water column mixing or stratification [Ramsing *et al.* 1996] may exert significant control over the spatial distribution of HABs and their timing. It has been well established, however, that the spatial distribution of planktonic organisms is highly heterogeneous [Pinel-Alloul 1995]. The phenomena is referred to as "planktonic patchiness" and is important when assessing the effects of HABs on overall water quality and ecosystem health. Planktonic patchiness occurs across a number of scales spanning from the microscopic to the size of oceans and is a function of both abiotic and biotic processes [Pinel-Alloul & Ghadouani 2007]. On smaller scales, biotic processes play the most important roles, while physical abiotic processes dominate with increasing scale. Pinel-Alloul & Ghadouani (2007) posit that these scales should be thought of in a hierarchical manner, that is, patchiness at smaller scales contributes to the structure of patchiness at larger scales. Studies are often constrained to a single scale with coarse sampling, for example a handful of sampling sites spread every few hundred meters across a reservoir as opposed to sites located every few meters; coarse scales may not

capture heterogeneity while smaller scales can be cost prohibitive. Physio-chemical conditions in the water column can also mirror this heterogeneity, with significant spatial differences in nutrient measurements and redox conditions [Barker *et al.* 2010] possibly driving HAB distribution.

Studies containing greater spatial and temporal detail in their sampling are needed to confirm the results of prior work and identify any relationships between T&O concentrations and observable environmental parameters that may have eluded previous studies. This study has generated a temporally and spatially expansive dataset of near-surface waters at Eagle Creek Reservoir (ECR) utilizing a suite of physical, chemical, and biological parameters to interrogate GSM or MIB production processes and if hotspots can be identified. We hypothesized that:

1. The spatial and temporal distributions of microbial cell, GSM, and MIB concentrations are controlled by hydrologic and chemical conditions.
2. Changes in nutrient levels and chemical conditions with the surface water control the formation and proliferation of HABs.

Methods:

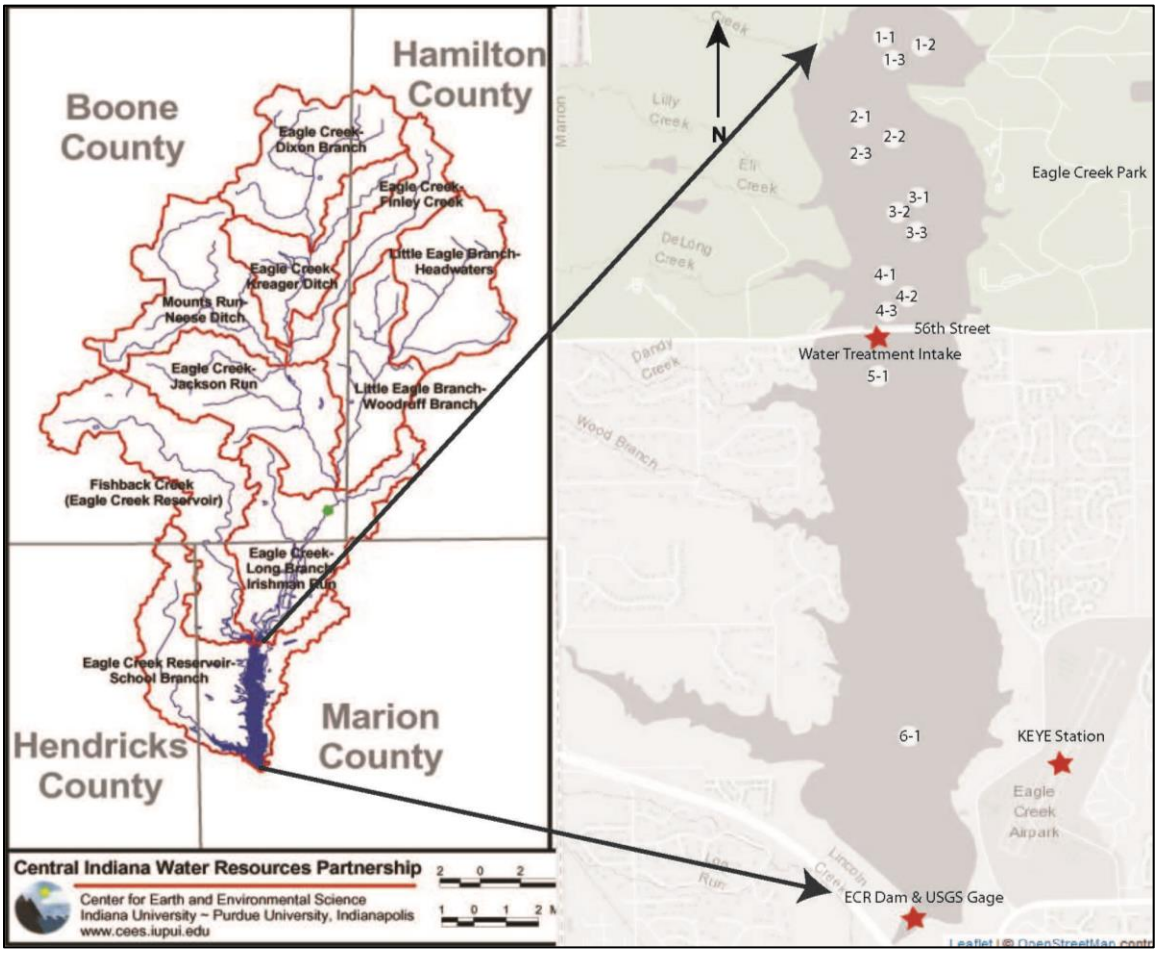

Figure 2.1: Overview of Eagle Creek Reservoir (ECR). Left pane modified from Tedesco *et al.* (2005) and illustrates the extent of ECR and its watershed. Right pane displays the reservoir itself, with sampling locations ("X-Y") and key locations (red stars) marked.

Study site

Eagle Creek Reservoir (39˚51'20"N, 86˚17'39"W), located in central Indiana, is an important freshwater resource for the city of Indianapolis that provides drinking water for over 80,000 citizens and serves as a recreational area for local residents [Pascual and Tedesco 2004]. The reservoir, like many freshwater bodies around the globe, is suffering from eutrophication driven by a large excess of nutrients within the watershed [Li *et al.*

2006]. Because of this, ECR has been experiencing annual algal blooms for well over a decade, with the first recorded measurement of T&O compounds dating back to at least 2000 [Pascual and Tedesco 2004].

The reservoir was constructed in 1969 as a means of flood control for the western portion of the city and later became a drinking water source as well. The watershed (HUC 05120201120) is substantially larger than the surface area of the reservoir at roughly 420 km^2 compared to 5.7 km^2. The land surface within the watershed is dominated by agriculture at almost 60% of the total area, with the remainder being urbanized/suburbanized land [Li *et al.* 2006]. Residence times of water within the reservoir average 50-60 days [Li *et al.* 2006] but are highly variable and can be much shorter during high-influx events (daily-averaged minimum of 0.87 days, maximum of 13,877 days during 2013) [Clercin and Druschel 2019]. Maximum depths reach 11 to 13 m in the southern basin nearest the dam but average 3 to 5 m in most other portions. ECR is a dimictic water body with seasonal stratification of the water column; mixings usually occur in April or May and October each year, while the water column becomes thermally stratified between June and September.

ECR has been the subject of several studies, a handful of which have been published, and investigations of T&O events at the reservoir are limited. One early study by Pascual and Tedesco (2004) examined physical and chemical conditions in the reservoir across four sampling sites as they relate to T&O concentrations and phytoplankton abundance before and after two copper algaecide treatments. It was found that the effectiveness of algaecide applications were tightly linked to bloom speciation and nutrient conditions within the reservoir, and algaecide was found to be most

appropriate for killing cyanobacteria when nutrient levels were primed for exponential growth. Clercin and Druschel (2019) sampled across multiple depths at a single location at ECR during 2013, with a focus on bacterial genetics and a detailed examination of bloom microorganisms' identities. Their work showed that several cyanobacteria species were likely responsible for GSM production while the actinobacteria *Streptomyces* was responsible for MIB. An algaecide treatment during the sampling period decimated the cyanobacterial community but had little effect on actinobacteria, and early season, mixed, cold, and nutrient-rich waters were associated with the growth of both communities.

Sampling regimen

Several hundred water samplings were performed during 2018 from the end of April through the middle of October, weekly, as weather permitted, across 14 locations in ECR (Figure 2.1). These locations were mostly focused on the northern basin of the reservoir where the watershed drains, with 12 locations spread across the northern area. The final two locations were next to the water treatment intake (at the northern most point of the southern basin) and in the southern, deeper, section of the reservoir, near the ECR dam. Samples were gathered off a small boat between 10AM and 1PM and immediately sub-sectioned for later analysis. Each day of sampling, water samples were collected using a Van Dorn sampler at 1m depth. Sub-samples included an unfiltered fraction (TP, Fe, Mn analyses) in a 50mL Falcon tube, another fraction (TDP, SRP, Ca, Mg, K, Na, DN analyses) filtered using a .22 µm GVWP filter in a 50mL Falcon tube, a third fraction of 50 mL spiked with ~1 mL of Lugol's Solution and stored in an amber glass vial for algae cell counts, another unfiltered fraction placed in an opaque 1L plastic bottle (Chl-a, PC, TSS analyses,) an unfiltered fraction placed in a 250mL amber bottle

with septum for T&O compound analysis, and finally a sample for genetic analysis collected by filtering 50 mL of reservoir water through a Sterivex filter. Sub-samples were immediately placed on ice in a cooler until they could be transported to the lab and refrigerated.

Hyperspectral remote sensing reflectance below the water surface (rrs) spectra were measured using a dual head Ocean Optics (OO) USB4000 system composed of two spectroradiometers (351-1047 nm, 0.2 nm spectral resolution, 3645 bands). One spectroradiometer was mounted on the end of a 2 m long pole pointed directly upwards, measuring the downwelling irradiance (Ed), while the other spectroradiometer equipped with a 25˚ field-of-view optical fiber was dipped approximately 5 cm below the water surface to measure below-surface upwelling radiance (Lu) at nadir. The final rrs were calculated following the procedure described by Gitelson *et al.* (2007).

Chemical and physical analyses

Metals analysis was performed using an inductively coupled plasma – optical emission spectrometry (ICP-OES) instrument [Perkin Elmer Optima 7000 DV OES] focused on Fe, Mn, Ca, Mg, K, and Na. Fe and Mn were analyzed using unfiltered samples, while Ca, Mg, K, and Na were analyzed using filtered samples following EPA method 200.7. Both unfiltered and filtered samples were spiked with HNO_3 to a concentration of approximately 2% and mixed well before experimentation. The analysis was carried out pursuant to the EPA method 200.7 for metals analysis in water.

Soluble reactive phosphorus (SRP) was determined by analyzing filtered samples, using the D'Angelo *et al.* (2001) method with the Malachite Green addendum modified

to work with microplate-sized sample volumes. Samples were run in quadruplicate using a 96-well microplate and BioTek Epoch 2 microplate reader [D'Angelo *et al.* 2001].

Total phosphorus (TP) and total dissolved phosphorus (TDP) were measured using the Hach Total Phosphate Reagent Set (#2742645) following method 8910. TP analysis was performed on unfiltered samples and TDP on their filtered counterparts, with absorbance measurements being made using the BioTek Epoch 2 microplate reader.

Dissolved nitrate (DN) was analyzed using filtered sub-samples and an ion chromatography system [Thermo Scientific Dionex ICS-1100 with AS-DV Autosampler, Dionex IonPac AS14 4x250 mm IC column and ASRS 300 4mm Suppressor, using 3.5mM Na_2CO_3/1mM $NaHCO_3$ eluent,] with samples ran in triplicate. Analysis was conducted following EPA method 300.00.

Algae cell counts were conducted by hand using microscopy with a Nagoette Bright Line Hemacytometer stage (Hausser Scientific, 0.500 mm stage depth) and an Olympus BX53 microscope (100x, 200x, and 400x magnifications) on unfiltered water samples treated with 1-2% Lugol's solution. Microorganisms were sorted and counted by phylum instead of attempting to identify individual species to facilitate higher sample throughput. The process followed the methods outlined by IUPUI's Center for Earth and Environmental Science (CEES) in their protocols for water monitoring [Clercin 2010].

Total Suspended Solids was measured using a simple gravimetric analysis, following EPA method 160.2. A GF/F filter was dry-ashed (103-105˚C) and weighed prior to filtering 300 mL of sample. The filter was then re-dried and weighed once more to determine the concentration of suspended solids in the original solution.

Chlorophyll A (Chl-*a*) content was measured following Arar (1997) and using a Shimadzu UV-2401PC UV-Vis Spectrophotometer. In short, 100 mL of sample was filtered using a GF/F filter and the filter placed in a Falcon tube containing 10 mL of 90% acetone and kept in a refrigerator overnight. The solution was collected and centrifuged, then the supernatant placed in a quartz cuvette for analysis [Arar 1997].

The phycocyanin (PC) analysis used is an amalgamation of several protocols [Siegelman *et al.* 1978, Sarada *et al.* 1999, Horváth *et al.* 2013]. 100 mL of sample was passed through a GF/F filter and the filter placed in a Falcon tube with 10 mL of a 0.1M phosphate buffer solution (10.5 g/L NaH_2PO_4, 4.4 g/L $Na_2HPO_4*12H_2O$ at pH 6). The sample was "freeze-thaw" treated three times (frozen for 2 hours, thawed, incubated in dark refrigeration for 24 h) before the solution was centrifuged. The supernatant was then collected and analyzed using a Turner Designs TD-700 Fluorometer to determine the concentration of PC.

The T&O compounds GSM and MIB were analyzed courtesy of the White River Water Treatment Plant labs of Citizens Energy Group. Samples were collected in 250 mL amber glass bottles with septum caps and delivered to the lab within an hour of field sampling concluding. Concentrations of GSM and MIB were quantified using Head-Space Solid-Phase Micro-Extraction (HS-SPME) coupled with a Gas Chromatography-Mass Spectrometry (GC-MS) instrument following method SM 6040D [APHA 2000].

Genetics data was measured in IUPUI's biology department using 16S qPCR assays as well as qPCR assays for MIB and GSM synthases on gathered water samples. Genetic material was attained from Sterivex filters collected at each sampling site and frozen for storage within several hours. The analysis followed the procedure detailed in

Koltsidou 2019, based partly on metagenomics data from ECR (Clercin 2019), utilizing primers developed for detecting MIB/GSM synthases of cyano- and actinobacterial organisms [Koltsidou 2019]. In short, DNA in the water samples was extracted using the DNeasy PowerWater Sterivex Kit from Qiagen. Lysis buffer was added to each Sterivex unit and mixed, then the lysate transferred to a bead beating tube for further lysis. Following this process DNA was captured on an MB Spin Column, washed, and the purified DNA eluted using a sterile elution buffer. This solution was immediately stored at -20˚C for later qPCR analyses. 16S assays as well as *geoA* and *MIBS* assays were conducted on aliquots of the purified DNA using a mix of TaqMan Universal PCR Master Mix, bovine serum albumin, and ECR-specific primers and probes designed for each experiment.

Data compilation

In addition to experimental and hyperspectral data, information about weather and hydrologic conditions was pooled together from online data repositories. Historical weather information, collected by the Eagle Creek Airpark KEYE station and the National Weather Service, was compiled and taken from *wunderground.com* for the months spanning the 2018 sampling season (data sourced from and consistent with National Weather Service station at the airpark). Hydrologic data was compiled using the USGS National Water Information System from gages 03353200 (“super”-gage located within the watershed) and 03353451 (located at the outflow of the dam). A substantial combined dataset (n = 333, variables = 132) containing information for every 2018 ECR sample was constructed.

Statistical analyses

The first analyses performed on the dataset were Pearson's (r) and Spearman's (r_s) correlation coefficients, which were intended as an exploratory technique to identify the direction, strength, and significance of pairwise relationships between any two variables. Following this, the 2018 data was log-transformed as well as centered and scaled before PCA and NMDS multivariate analyses were carried out to reduce the number of dimensions in the dataset and identify any variables with significant control over dataset variability. Finally the data was fed into the "dismo" and "gbm" [Elith *et al.* 2008] R packages to generate a Boosted Regression Tree (BRT) model capable of predicting MIB concentrations, utilizing the power of machine-learning techniques.

Data imputation

A major problem with large datasets collected from the environment is missing or incomplete data due to any number of reasons. Analyses like pairwise correlations can easily overcome modest amounts of missing information by ignoring offending entries during each pairwise test, leaving subsequent tests unaffected. PCA and NMDS necessitate that every sample has information for every variable, which can quickly result in some variables and many samples being removed from the dataset. In the case of the ECR data entire sampling days are lost, shrinking the number of observations considerably ($n < 200$). Missing data can be imputed through simple methods like using the mean or median of a variable to fill in missing observations, but this can heavily skew the data and distort the results of PCA or NMDS.

More sophisticated imputation methods exist, each with their own strengths and weaknesses. For this study the "missMDA" package for R [Josse and Husson 2016] was

used for PCA as well as data imputation. A comparative study found that the imputation algorithm used by missMDA, termed iterative PCA (IPCA), was far superior to mean/median imputation and surpassed many other advanced methods across a number of scenarios [Dray and Josse 2015]. Furthermore, the ECR dataset was analyzed both with and without imputation using PCA and NMDS and the results were very similar, but the imputed dataset gives more precision and allows the model to be extended over greater time and space (see Figures A1.1 and A1.2 for a comparison between non-imputed and imputed results).

Results:

All of the collected data was compiled into a large dataset consisting of 333 observations and 115 measured variables ("OaXX" reflectance bands (see Appendix X, Sheet 1: "Notes" for further detail) were removed from analyses after no significant correlation with environmental variables was found). The full dataset is available in the supplementary materials (Appendix X) including summary statistics and notes for each variable, information about measurement units, and monthly weather trends during the sampling period. A subsection of the dataset, containing the more prominent variables seen in multiple analyses, was Log10 transformed and graphed as box-and-whisker plots (Figure 2.2). The data was also centered and scaled (Figure B2.1) revealing that a majority of variables contain outliers, are positively skewed, and non-normal. Outliers were not removed during statistical analyses, as these extreme values are likely event-related and contain important information about ECR processes.

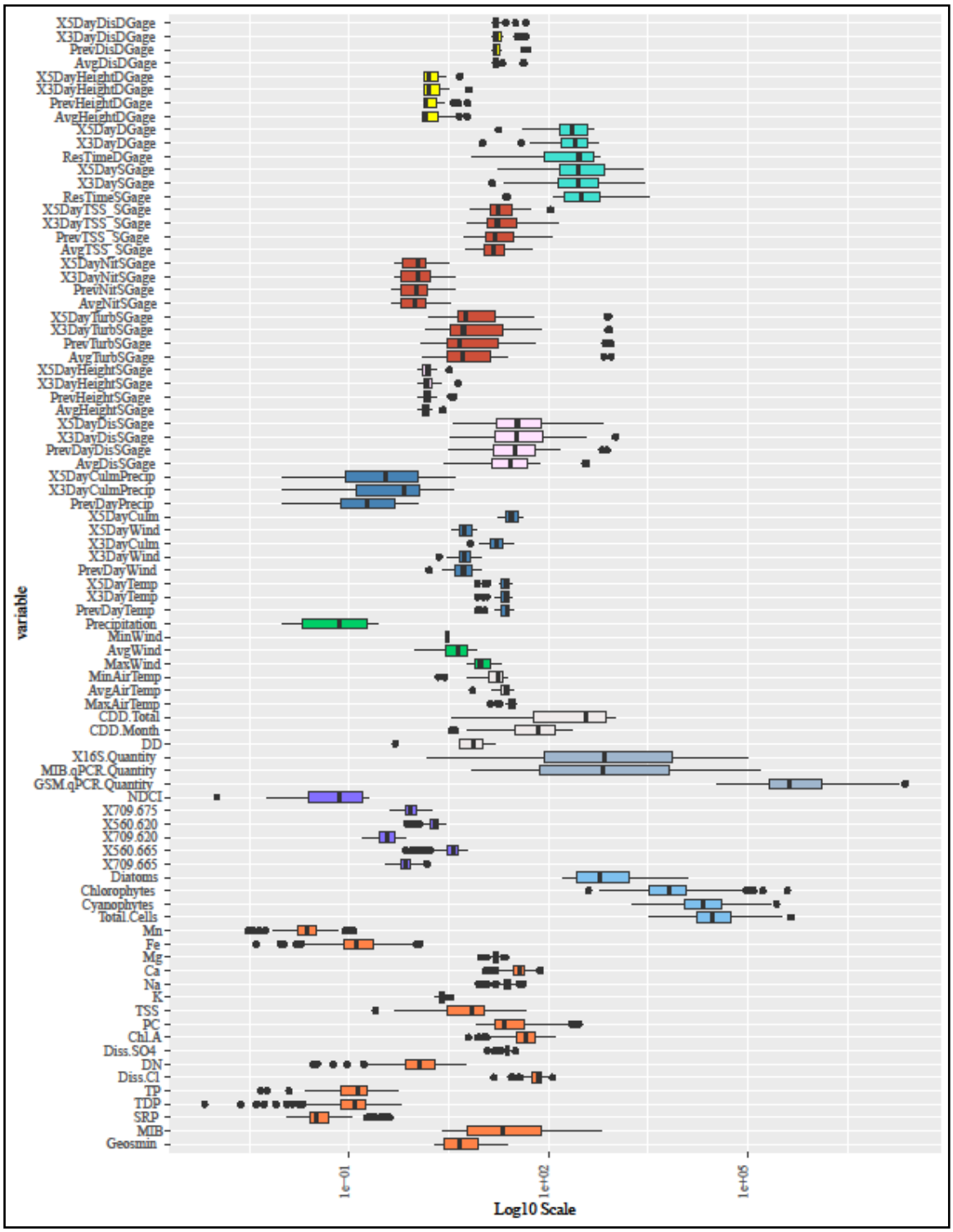

Figure 2.2: Box-and-whisker plots for a large subset of the variables, showing differences in magnitude and spread. Outliers are indicated by black dots on the graph. Data has been Log10 transformed. A centered and scaled version is also available (Figure B2.1).

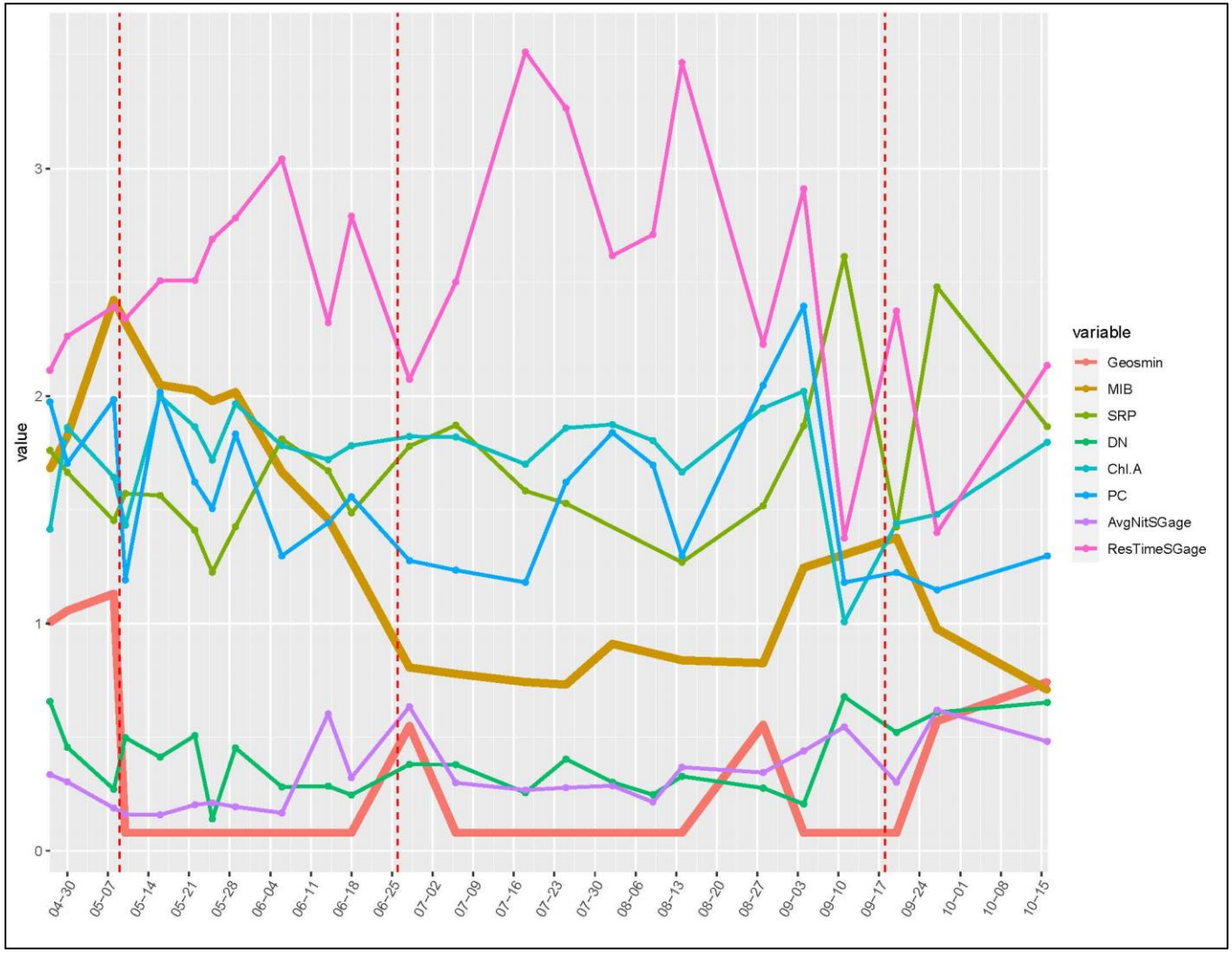

Figure 2.3: Timeseries plot showing changes in 8 different variables at sampling site “1-1” across 2018. Values have been Log10 scaled to fit on the graph. Red dotted lines indicate dates of copper algaecide

MIB concentrations in 2018 were highly variable and exceeded 600 ng/mL at some locations (Figure 2.3). Concentrations of both MIB and GSM peaked during early May before algaecide was applied to the reservoir, which had an immediate effect on GSM levels and algal biomass (Chl-*a* & PC) but did little to change the MIB concentration (Figure 2.3). Chl-*a* and PC measurements quickly rebounded within that same month, and MIB concentrations began a slow decline through June. GSM began rising in concentration again during late June, coinciding with decreased residence times and an influx of N from the watershed. A second algaecide treatment was applied at the end of June, though the effect on GSM was delayed several weeks and also coincided with increasing residence times and decreased nutrient input (Figure 2.3). MIB, once again, was mostly unaffected and began to slowly rise in July. Another peak in GSM occurred before September alongside increased Chl-*a* and PC concentrations, and then crashed as dramatic shifts in residence time were occurring. The third and final algaecide treatment of the year occurred shortly after, in mid-September, and was promptly followed by sharp increases in GSM, Chl-*a*, and PC.

Attempts were made to display spatial differences across the reservoir in the form of a heat map using kriging techniques but were largely unsuccessful. The sill of a semivariogram showed that interpolation of surface waters beyond several feet was ill-advised and would lead to large inaccuracies. Spatial variation was instead visualized using “bubble maps” of sampling points, each colored according to their corresponding measurement ranges for each variable. Examples of these plots are available for SRP and MIB in Figures B2.2 and B2.3 and the R script in Appendix Y includes the necessary tools to generate these graphs for any continuous variable in the dataset. Overall there

were no obvious spatial differences for the examined variables and no T&O or HAB hotspots were detected. These graphs do show clear temporal differences however, with major changes within the reservoir occurring across time rather than space.

Pearson's and Spearman's correlation coefficients were calculated for all pairwise variable combinations and any insignificant ($p > 0.05$) results removed (Figures 2.4 and B2.4). Calculated correlations were plotted graphically and important variable interactions identified in the dataset based on high r or r_s coefficients. Both tests produced similar correlations between many of the variable combinations, but there were stark differences in strength and/or significance for several important relationships. Figure 2.4 represents the results of Spearman's rank correlation, and a similar figure for the results of Pearson's correlation is available as supplementary material (Figure B2.4). In Figure 2.4, the strongest correlations with MIB concentration are for MIB qPCR counts ($r_s = 0.75$, $p = 7.8 \times 10^{-22}$) and the cumulative number of degree days ($r_s = -0.68$, $p = 1.01 \times 10^{-13}$). Weaker correlations include TSS ($r_s = -0.57$, $p = 1.72 \times 10^{-11}$), nitrate measured at the watershed gage ($r_s = -0.52$, $p = 5.12 \times 10^{-7}$), GSM qPCR counts ($r_s = 0.48$, $p = 3.61 \times 10^{-9}$), SRP ($r_s = -0.45$, $p = 1.79 \times 10^{-5}$), and dissolved chloride ($r_s = -0.44$, $p = 3.70 \times 10^{-3}$). GSM was most strongly associated with calcium ($r_s = 0.62$, $p = 9.45 \times 10^{-11}$), five day average wind speed ($r_s = 0.55$, $p = 6.19 \times 10^{-41}$), average discharge within the watershed ($r_s = 0.54$, $p = 9.00 \times 10^{-38}$), minimum air temperature ($r_s = -0.47$, $p = 2.71 \times 10^{-35}$), and dissolved nitrate ($r_s = 0.44$, $p = 5.90 \times 10^{-40}$). DN:TDP ratios were also included in the dataset but showed no significant correlations with either T&O compound.

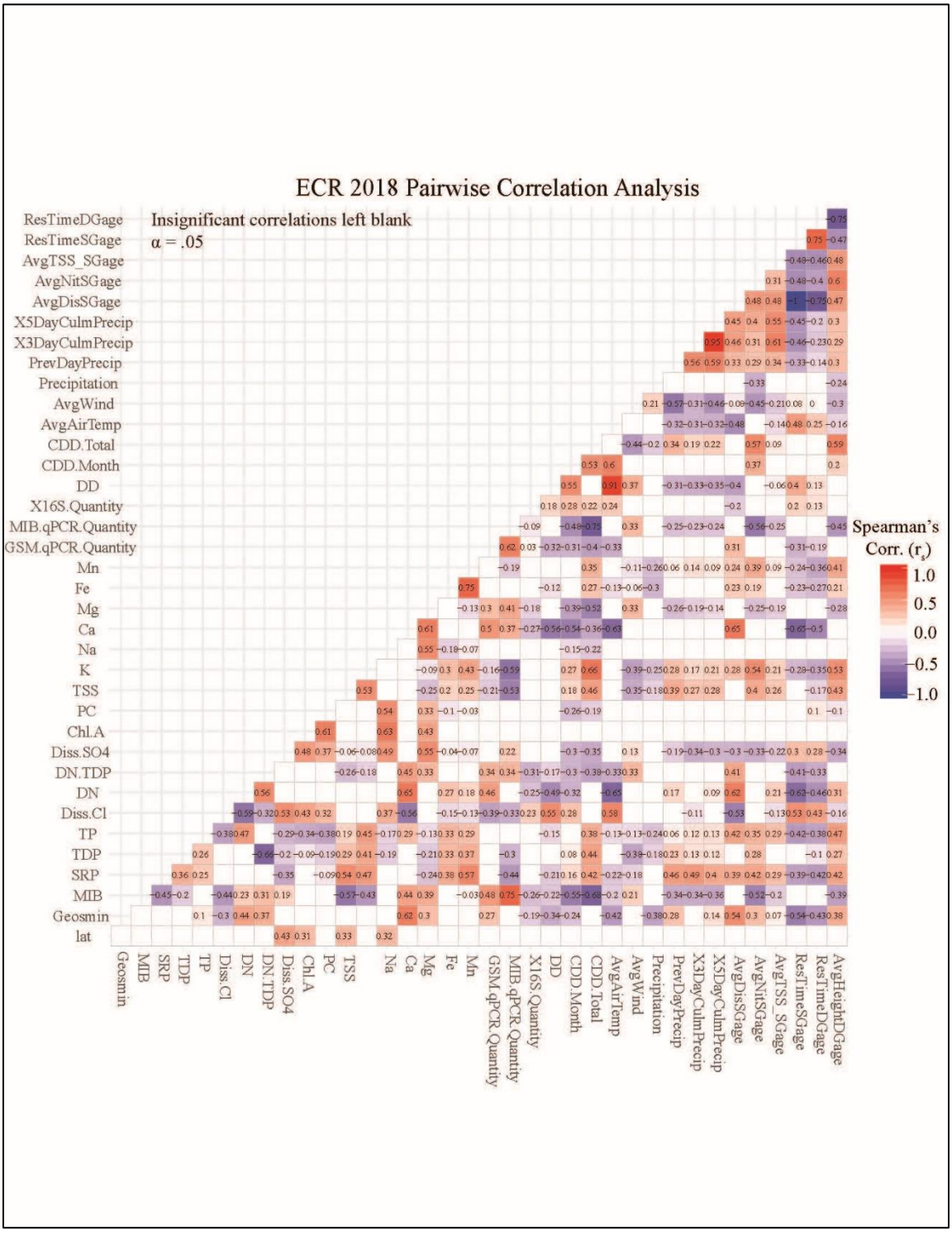

Figure 2.4: Spearman's Correlation matrix for a large selection of the dataset. Correlations with a p-value greater than 0.05 are left blank. The data and R script needed to re-create this and other matrices are available as part of the Appendix X and Y.

PCA was performed on the imputed dataset after it was log10 transformed, centered, and scaled to avoid any influences from variable scaling or non-normal variable distributions (Figure B2.5). The calculated principal components (PC'X') described approximately 47% of the data's variance within its first two dimensions, roughly 56% in the first three, and slightly more than 70% if five dimensions were included. The R script used to generate these graphics is available in Appendix Y and includes tools for viewing individual component loadings. Ideally, 70% or more of variance in a dataset would be explained by the first two or three principal components [McKillup and Dyar 2010], but given the size and complexity of the ECR dataset 56% is a considerably strong result. PC1 is predominantly controlled by discharge measured at both gages and SRP in the positive direction, and residence time, dissolved chloride, and temperature in the negative direction. PC2 is mostly composed of temperature, degree day measures, and universal primer qPCR counts in the positive direction and GSM, MIB, wind, turbidity, cell counts, and GSM/MIB qPCR counts in the opposite. When looking at the sample biplot (Figure B2.5) there is a clear cyclical trend throughout the season with conditions during April most closely matching those during October after tracing out a path through the summer months, resembling a system with hysteresis.

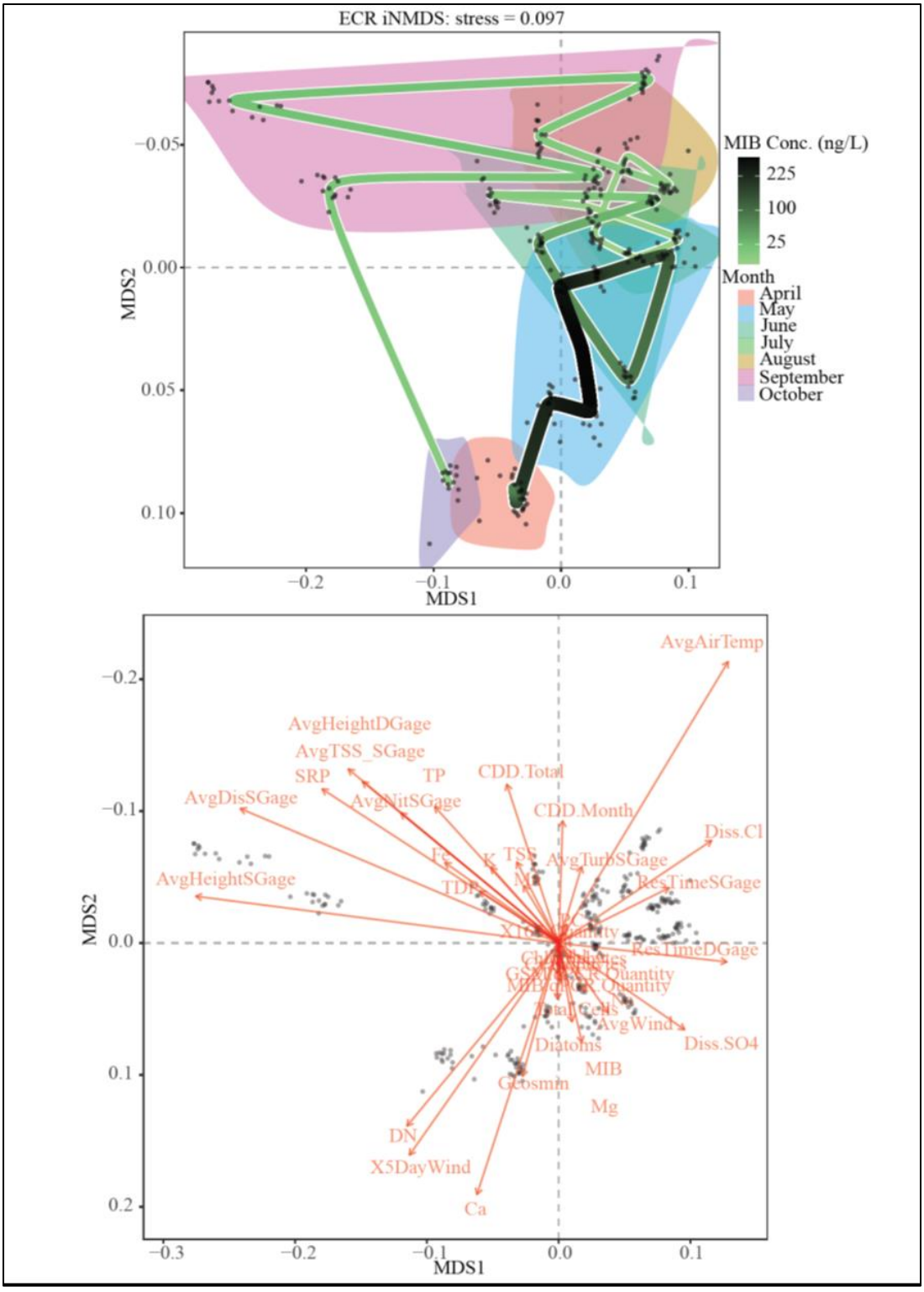

Figure 2.5: Results of NMDS analysis on the imputed dataset ("iNMDS"). Sample biplot (top) with sampling months encircled by polygons, line is traced through the average of all samples, by date from the start of season to the end. Simplified NMDS vector space

NMDS analysis was carried out on the same imputed dataset after it had been centered/scaled and the data shifted positively, by variable, to eliminate any negative values that can interfere with computations (Figure 2.5). NMDS performed exceptionally well by metric standards with a stress less than 0.1, and seemed better able to group similar variables together, but provided much of the same information as PCA. In fact, a NMDS biplot for the entire season is strikingly consistent to the PCA biplot (Figures 2.5 and B2.5). The stress of fitted values for individual samples in the NMDS model can be examined using the provided R script (Appendix Y), and, shows that the overall stress is likely closer to 0.05 than 0.1 if a few outlier samples were removed. MDS1 is largely associated with discharge and nutrient measures while MDS2 is related to temperature, wind, genetic/cell counts, and T&O concentration. In general, the first dimension can be thought of as a discharge and nutrient load axis, while the second as a mixing, stratification, and biologic axis. While NMDS does not provide variable loadings for each axis like PCA, it can still generate a list of variable 'vectors' for the NMDS axes being plotted. The position of the variables in these plots is often thought of as the 'maxima' for that variable, and when interpreted in this way the NMDS results are further consistent with those of PCA.

The raw dataset was fed into the "dismo" and "gbm" R packages to construct a Boosted Regression Tree (BRT) model with MIB concentration as the response variable. MIB concentration was square-root transformed prior to this in order to bring the variable's distribution closer to Gaussian (required for continuous response variables in BRT applications). The constructed BRT model consisted of 2900 trees with a mean residual deviance of 0.427 and an estimated 10-fold cross-validation deviance of 1.596

(Figure 2.6). The training data correlation was 0.991 with a 10-fold cross-validation correlation of 0.969, but without a similarly sized dataset from a separate ECR season it was impossible to test the model's predictive capabilities on independent data.

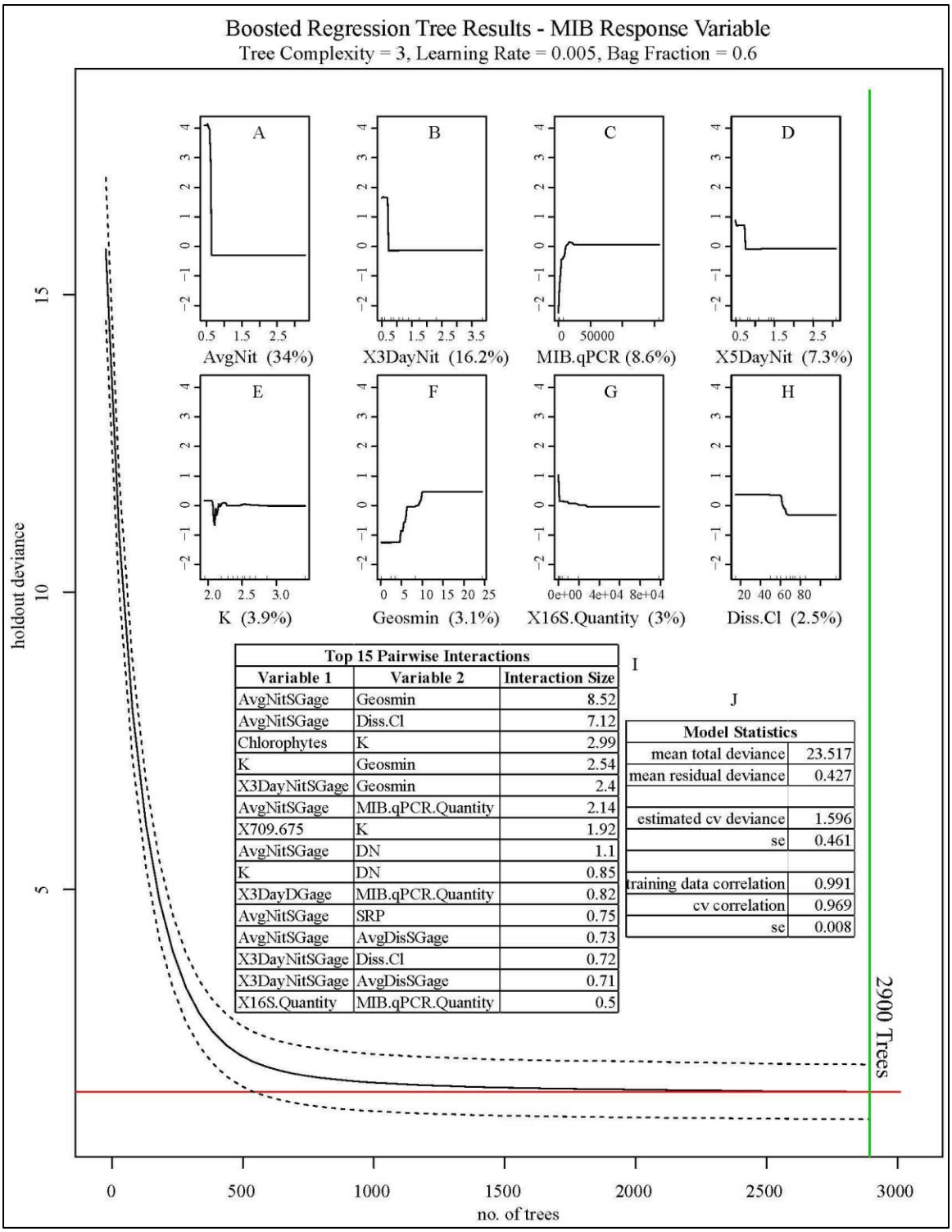

Figure 2.6: Summarized results of the BRT model. Fitted functions for all predictors with relative contributions > 2% shown, as well as the top 15 pairwise interactions in the model.

The single strongest predictor of MIB concentration during 2018, based on BRT results, was overwhelmingly the nitrate concentration measured at the USGS gaging station within the watershed, followed by GSM and dissolved chloride concentrations. The fitted functions for each predictor show that low nitrate levels correspond to the highest MIB concentrations (Figure 2.6 A,B,D), with varying predictive strength depending on the time-frame of nitrate measurement (day-of, prior 3-day average, prior 5-day average). In addition, high GSM levels and low dissolved chloride were associated with high MIB. The R packages used to create the model also include basic tools for identifying pairwise interactions in the final BRT model, of which the interactions between nitrate concentration and both GSM and dissolved chloride were by far the strongest (Figure 2.6 I).

Discussion:

The primary goal of this study is to find links between T&O compound concentrations and other measurable environmental parameters. Relationships between GSM/MIB and other variables can be elucidated through statistical analyses with the intent of better understanding HAB processes and T&O event dynamics. Knowledge of these processes will hopefully inform more efficient and successful management or remediation strategies for freshwater systems in the future. However, the following discussion will focus primarily on MIB and its links with environmental parameters as MIB was by far the T&O compound of most concern for 2018. GSM rarely exceeded 10 ng/L and peaked at roughly 25 ng/L during May, while MIB was consistently in excess of 20 ng/L and peaked well above 600 ng/L during May (Figure 2.3). Aside from assessing

HAB and T&O event processes through variable relationships, we also seek to identify any temporal or spatial trends/patterns in the dataset which may further enhance our understanding of these relationships.

Physical-chemical-biological links to T&O production

The 2018 dataset contains a wealth of physical, chemical, biological, and spectral information that was interrogated to find any links between T&O production and any available variables. Several interesting associations between MIB qPCR counts (MIB synthase targeted primer) and other parameters quickly became apparent. A PCA model of the data (Figure B2.5) was able to successfully group GSM and MIB with their respective qPCR counts while separating the two groups from one another, indicating that the qPCR techniques developed for ECR microorganisms (Koltsidou 2019) are able to differentiate T&O compound production by actinobacteria versus cyanobacteria. Additionally the PCA model (Figure B2.5) linked microscopy total cell counts, but not Chl-*a* or PC, to MIB concentrations and MIB qPCR counts. The NMDS model (Figure 2.5) did not display either phenomenon, with Chl-*a*, PC, and cell counts only having relatively short vectors pointing in a similar direction to GSM/MIB. The BRT model (Figure 2.6) identified MIB qPCR as the third strongest predictor of MIB concentrations, and the fitted function shows that low MIB qPCR counts have a strong positive relationship with MIB levels. Past a certain level of MIB qPCR counts, however, the BRT function (Figure 2.6 C) quickly flat lines and this relationship ceases to exist. This, coupled with the fact that Chl-*a* and PC were not significant variables in the BRT, PCA, or NMDS models (Figures 2.6, B2.5, 2.5), suggests that the number of microorganisms or overall bloom biomass is largely unimportant at ECR for predicting T&O production.

Data suggests that even a relatively small number of microorganisms are capable of producing substantial T&O concentrations. Here, it should be noted that microscopy cell counts (total cells, cyanophytes, chlorophytes, and diatoms) also showed no significant relationship to GSM/MIB.

Although GSM concentrations remained relatively low throughout 2018 when compared to MIB, the outcomes here strongly hint at a co-occurrence of GSM and MIB or some kind of previously unknown relationship between the two compounds and/or their progenitors. Cross-correlation analyses (Figures 2.4 and B2.4) found a strong correlation between MIB concentration and MIBS qPCR ($r_s = 0.75$, $p = 6.45x10^{-31}$). Interestingly, GSM was the most correlated with MIB ($r = 0.30$, $p = 4.56x10^{-5}$) rather than *geo-A* qPCR ($r_s = 0.27$, $p = 3.10x10^{-6}$), and MIB had a stronger correlation with *geo-A* qPCR (r_s 0.48, $p = 2.08x10^{-11}$) than GSM did. This could be an artifact from the magnitude difference between GSM and MIB concentrations during the season, with the larger concentrations and variability of MIB driving much of the association. Further evidence of a relationship between GSM and MIB comes from the PCA and NMDS models (Figures B2.5 and 2.5) where the two molecules share roughly the same dimension. Furthermore, the BRT model (Figure 2.6 F) illustrated that GSM in an important predictor of MIB concentrations. Given the many linkages between GSM and MIB one might assume that the two compounds are being produced by same microorganism or group of microorganisms, but the findings of Clercin and Druschel (2019) show that this is likely not the case. The involvement of separate organisms is further supported by the separation/differentiation of the GSM and MIB qPCR counts observed in the PCA model (Figure B2.5). The organisms responsible for GSM/MIB

likely occupy different environmental niches and certainly respond in different ways to algaecide treatments (Pascual and Tedesco 2004), making it difficult to argue that T&O co-occurrence is caused by factors such as bloom growth/death. It is more plausible then that these two compounds are used for communication, either as part of an inter-species signaling chain involving both molecules or as two similar intra-community signaling processes responding to the same environmental changes at ECR.

Chl-*a* (Vörös and Padisák 1991) and PC (Ahn *et al.* 2007) concentrations have been shown to be effective proxies for estimating phytoplankton and cyanobacterial bloom mass. Chl-*a* and PC were not associated with T&O concentrations at ECR during 2018 (Figures 2.5 and B2.5). This is consistent with previous observations by Clercin and Druschel (2019) that found neither parameter was associated with high T&O concentrations, but rather were tied to late season blooms which did not generate significant secondary metabolites (T&O or neurotoxins). This is also evident graphically in the 2018 data (Figure 2.3), as spikes and dips in Chl-*a*/PC are not correlated with changes in MIB throughout the year. GSM is seen to drop dramatically with the season's first algaecide treatment in early May and a sudden dip in Chl-*a*/PC, though Chl-*a* and PC quickly rebounded while GSM did not (Figure 2.3). Throughout the remainder of the year GSM rises during or shortly after sudden increases in PC but does not approach the odor threshold and remains relatively low. MIB concentrations throughout time were on a steady decline following the first algaecide application and did not respond to fluctuations in Chl-*a*/PC until the very end of the season, where a rise in both parameters occurs during a steep decline in MIB concentrations below the odor threshold. Based on the lack of association between Chl-*a*/PC and the two T&O compounds, as well as the

timing of events, one can conclude that elevated T&O concentrations at ECR are not linked to large populations of producers and the result of a bloom die-off and decomposition as seen in some other systems [Ma *et al.* 2013]. If this was indeed the case at ECR it would be expected that major declines in Chl-*a*/PC measurements would precede sharp increases in T&O concentrations, which is clearly not the case (Figure 2.3).

The only nutrients with moderate (r or $r_s > 0.4$, $p < 0.05$) correlations to T&O concentrations in the dataset were different measurements of N, both in the reservoir and within the watershed, and SRP (Figures 2.5 and B2.5). NMDS (Figure 2.5) results associated T&O compounds in roughly similar vector space with DN, but closer to orthogonal with TP, TDP, SRP, and nitrate in the watershed (AvgNitSGage). The PCA (Figure B2.5) vector space was similar, but DN was much more closely associated with GSM and nearing orthogonal with respect to MIB. The BRT model (Figure 2.6) identified nitrate in the ECR watershed, but not within the reservoir (DN), as the most important predictor of MIB concentrations with P measurements being much less important. The fitted functions for watershed nitrate in the BRT model show that the lowest concentrations of nitrate led to the highest predicted MIB concentrations, with a quickly lessening effect as nitrate concentrations increased. As nitrate concentrations decrease to near 0.5 ppm the fitted function flat lines slightly below zero on the y-axis, meaning concentrations beyond that point contribute only slightly to lower predicted MIB concentrations. This may reflect the existence of a threshold at which N becomes limiting. N and P limitation are often thought of jointly, with an excess of one leading to limitation by the other, and viewed in the context of TN:TP ratios. The Redfield ratio (N:P = 16 mol:mol) is often employed where values below 16 are indicative of N-

limitation and values above indicating P-limitation [Andersen et al 2020]. DN:TDP ratios during the 2018 sampling season ranged from roughly 0.1 mol:mol to 814 mol:mol, with an average of approximately 22 and standard deviation of 67. This ratio remained above the 16 mol:mol threshold throughout most of the season, approaching that limit during the early-May T&O event and dropping below it briefly afterwards. The threshold at which N forms become the limiting nutrient suggests that, at least for some systems, the concentration of N needed to induce limitation is so appreciably low (7.29 µmol/L NO_3 + NH_4) as to rule out N reduction as a practical remediation strategy [Chaffin *et al.* 2014]. The trend of decreased nitrogen availability occurring before/during a T&O event is confirmed by the time series plot (Figure 2.3) where peaks in MIB occur during troughs in watershed nitrate concentrations, and the Spearman's correlation matrix (Figure 2.4) where MIB is negatively correlated with AvgNitSGage ($r = -0.52$, $p < 0.05$). The evidence points towards some process in which changes in available nitrogen trigger increasing or decreasing MIB concentrations at ECR. There are likely two general interpretations of what this process might be:

1. Given the importance of N and P to phytoplankton growth [Lean *et al.* 1987, Xu et al 2010, Ma *et al.* 2013] a sudden limitation of nitrogen, either through a change in total available nitrogen or change in nitrogen speciation, triggers a bloom die-off. Subsequently, the decomposing microbes release a significant flux of T&O compounds (in this case MIB) into the overlying water column.
2. Alternatively, significant cell death from N limitation and decomposition is not occurring or is not the cause of elevated T&O concentrations. Changes in

MIB concentration are therefore a response by members of the algal community to changing nutrient conditions.

If bloom die-off and decomposition were the predominant process we would expect to see a series of observable events, from a sudden limitation in nutrients to algal death and then a T&O event. Instead, changes in T&O concentrations are a response by microbes to changing nutrient conditions similar to the findings of Perkins *et al.* (2019). In the timeline shown in Figure 2.3 peak MIB concentrations during the year actually preceded an algaecide treatment and significant drop in algal biomass, which does not support the first hypothesis. During this same period both DN and watershed nitrate were steadily decreasing as MIB rose to its peak, and a sudden increase in DN coincided with a shift to falling MIB levels. Additionally, MIB concentrations continued to slowly decline even after an abrupt rebound in algal biomass. It is far more plausible that this major, early season T&O event was the result of MIB producers responding to changing nutrient availability, specifically the degree or state of N limitation, rather than decaying en masse and releasing their contents.

Seasonal variations in ECR system

Watershed and dam gage discharge measurements are important in explaining differences in the conditions at ECR across the season. Gage discharge and gage height at both the watershed and dam locations significantly contribute to the major axis of the PCA (Figure B2.5) and NMDS (Figure 2.5) models. Gage discharge measurements at the watershed show significant (p-values less than 0.05) linear correlations with TP ($r = 0.59$), TDP ($r = 0.48$), SRP ($r = 0.88$), and nitrate gage measurements ($r = 0.61$) (Figure B2.4) as well as monotonic correlations with DN measured in the reservoir ($r_s = 0.62$)

(Figure 2.4). These strong correlations, in addition to the contributions of discharge to explaining PCA/NMDS model variability, suggest that nutrient conditions at ECR are significantly controlled by the state of the watershed. Figure 2.3 traces residence time (calculated using daily averages of super gage discharge and assuming constant reservoir volume) throughout 2018, and a general trend of increasing residence time from April to July followed by sharp declines in August and September is observed. Residence time during 2018 was highly variable, maximum residence times reached 3,256 days while minimums were as low as 22.7 days, averaging 634.2 days with a standard deviation of 838.7 days (Appendix X). When estimated using the outflow of the ECR dam (measured as discharge at the gaging station) residence times averaged 286 days with a standard deviation of 193.7 days and a max/min of 594.4/7.2 days (Appendix X). ECR clearly experienced periods of lake-like conditions (high residence times) and more riverine-like conditions (low residence times) during the year, which may have contributed to nutrient conditions in the water column. Under lake-like conditions redox stratification is more probable, and stratification has been observed at ECR in the past [Clercin and Druschel 2019]. Studies have demonstrated that, for shallow lakes, redox stratification and internal loading can significantly contribute to nutrient conditions [Petticrew and Arocena 2001, Søndergaard *et al.* 2003, Smith *et al.* 2011] though no studies of ECR sediment dynamics under changing redox conditions exist so it is not possible to assess any involvement of internal nutrient loading.

MIB vs. Discharge			GSM vs. Discharge		
Lag (days)	**Correlation**	***p***	**Lag (days)**	**Correlation**	***p***
-1	-0.195	3.62E-01	-1	0.017	9.37E-01
-2	-0.170	4.26E-01	-2	0.048	8.25E-01
-3	-0.130	5.44E-01	-3	0.285	1.76E-01
-4	-0.057	7.92E-01	-4	0.304	1.48E-01
-5	0.004	9.83E-01	-5	0.326	1.20E-01
…	…	…	…	…	…
-22	0.403	5.12E-02	-22	0.633	1.16E-03
-23	0.121	5.73E-01	-23	0.604	2.16E-03
-24	0.094	6.64E-01	-24	0.545	6.62E-03
-25	0.231	2.77E-01	-25	0.473	2.07E-02
-26	0.061	7.77E-01	-26	0.541	7.13E-03
-27	0.084	6.95E-01	-27	0.579	4.44E-03
-28	0.108	6.15E-01	-28	0.558	6.41E-03
-29	0.216	3.10E-01	-29	0.583	4.12E-03
-30	0.319	1.29E-01	-30	0.634	1.94E-03
-31	0.282	1.82E-01	-31	0.609	3.18E-03
-32	0.386	6.28E-02	-32	0.543	1.01E-02
-33	0.531	7.65E-03	-33	0.517	1.50E-02
-34	0.715	8.70E-05	-34	0.595	4.18E-03
-35	0.768	1.20E-05	-35	0.512	1.60E-02
-36	0.537	6.81E-03	-36	0.452	3.59E-02
-37	0.458	2.46E-02	-37	0.414	5.67E-02
-38	0.307	1.45E-01	-38	0.416	6.22E-02
-39	0.794	3.68E-06	-39	0.509	1.98E-02
-40	0.906	1.10E-09	-40	0.520	2.01E-02
-41	0.850	1.47E-07	-41	0.481	3.33E-02
…	…	…	…	…	…

Table 2.1: Lists the Pearson Correlation Coefficient and p-value for MIB and GSM vs. discharge measured within the watershed. When lag time between peak discharge and peak T&O concentrations is accounted for significant correlations are found.

It has long been recognized that bacterial communities undergo several different phases throughout their growth, the first of which is commonly referred to as lag phase [Penfold 1914]. During lag phase the bacteria are effectively dormant, resulting in a lag time between the introduction of nutrients to a system and exponential growth of the community. Though the process is not fully understood it is believed that lag phase

represents a time of cell maturation and adaptation to new nutrient conditions and is an important consideration when modeling bacterial growth [Rolfe *et al.* 2011, Schultz and Kishony 2013]. It is reasonable, then, to assume that this process occurs for bacteria and microbes within algal blooms as well. Several studies have successfully drawn connections between HAB dynamics and biological, physical, chemical, or hydrological parameters when consideration of a lag phase was included. Harred and Campbell (2014) found strong positive correlations between two planktonic organisms when time lag was included, highlighting the use of leading indicators for predicting future blooms. Sarkar *et al.* (2007) and Chattopadhyay *et al.* (2002) were able to construct informative mathematical models of phytoplankton-zooplankton interactions when the lag associated with growth and maturation of different organisms was accounted for. Clercin and Druschel (2019) noted strong correlations between watershed discharge and both T&O compounds at ECR when adjusted for lag time, thought to be driven by the delayed bloom growth following nutrient pulses. This was further investigated here, MIB and GSM were averaged across all sampling sites by date, and then Pearson's cross-correlation performed on a combined T&O and discharge matrix. The process was done iteratively, adjusting the discharge data by increments of one day and re-running the analysis. Table 2.1 displays the results, and strong, significant correlations between both T&O compounds and discharge were observed. MIB's strongest correlation with discharge was found at 40 days of lag ($r = 0.906$, $p = 1.10 \times 10^{-9}$) while GSM's occurred with 30 days of lag ($r = 0.634$, $p = 1.94 \times 10^{-3}$). This is strong evidence that an early season nutrient pulse, most likely in April or late March, is driving peak T&O concentrations in May.

Seasonal weather patterns also contribute to the state of ECR, evidenced by the importance of air temperature and wind speeds in the PCA and NMDS models (Figures B2.5 and 2.5) and their contribution to the observed hysteresis pattern. Information on weather conditions is available as a part of Appendix X, but is summarized here. April was by far the coldest and windiest month during the 2018 sampling period, and saw nearly 10 cm of precipitation. The time span from June through August was the hottest and least windy of the season, with June receiving over 15 cm of precipitation. September was the second rainiest month behind June (over 13 cm of precipitation), was slightly colder than May had been, and was the least windy month. Pearson and Spearman correlations only found weak to moderate associations between GSM/MIB and weather variables, but with significantly stronger associations between these same weather variables and important nutrients. GSM was correlated to air temperature ($r = -0.52$, $p = 1.29x10^{-22}$), averaged five day wind speeds ($r = 0.47$, $p = 7.87x10^{-22}$), and precipitation on the previous day ($r_s = 0.28$, $p = 3.91x10^{-14}$). MIB was correlated with previous day's precipitation ($r_s = -0.34$, $p = 1.48x10^{-6}$), average wind ($r_s = 0.21$, $p = 4.76x10^{-9}$), and average air temperature ($r_s = -0.20$, $p = 1.95x10^{-3}$). Weather variables were much more closely linked with nutrients with significant relationships found between SRP and 3 day cumulative precipitation ($r = 0.74$, $p = 2.76x10^{-50}$), 3 day average wind speeds ($r = 0.44$, $p = 6.14x10^{-20}$), and average temperature ($r = -0.33$, $p = 2.03x10^{-14}$). Total Phosphorus (TP) and Total Dissolved Phosphorus (TDP) had weaker but similar correlations to these same variables. Nitrate measured within the reservoir (DN) was strongly correlated with average temperature ($r_s = -0.65$, $p = 1.35x10^{-50}$) and 5 day average wind speeds ($r_s = 0.56$,

$p = 1.43x10^{-39}$) while watershed nitrate was correlated with average wind speed (r_s = -0.45, $p = 2.51x10^{-24}$) and previous day's precipitation ($r = 0.58$, $p = 1.50x10^{-33}$).

Alternative stable states are a common model for interpreting community dynamics in ecology, and hysteresis is oftentimes invoked as a necessary characteristic [Beisner *et al.* 2003]. Alternative stable states posits that a system contains a number of states which are stable and resistant to perturbation. Disruptions to the system eventually rebound back to one of these stable states, but large enough changes can permanently move the system out of one stable state and into another. Hysteresis is the idea that parameter changes lead to shifts in equilibrium position (stable state), and the state of a system/community is dependent on its previous history. An example involving a lake ecosystem would be low-P and high-P states driven by the intensity of farming in the catchment and flushing within the lake. As the proportion of intensely fertilized farms versus conservatively tilled farms increases the amount of P entering the lake increases. At a certain point this increase triggers widespread algal blooms (high-P state), and this transition threshold is lowered with decreasing efflux from the lake. The threshold for reverse transition is significantly different, with a much lower proportion of intensive farms needed to transition back to a new low-P state than were needed to leave the original low-P state, exemplifying the idea of hysteresis. The difference in intensive farm proportions needed to trigger these transitions are most exacerbated when lake efflux is low [Dent *et al.* 2002]. In the NMDS model biplot (Figure 2.5) a path is traced through different sampling dates and months, showing a hysteretic seasonal pattern. The x-axis is largely defined as a discharge and nutrient load dimension, with increased discharge and nutrients in the negative direction and increasing residence times in the positive direction.

The y-axis is largely defined as a dimension with warm, stratified waters in the negative direction and cold, well mixed waters in the positive direction. Starting in April, reservoir conditions plot towards the bottom left and are characterized by elevated discharge, high nutrient inputs, cold temperatures, and well mixed waters. T&O concentrations peak in May when the water column is still mixed, but discharge and nutrient loads have slowed down (Figure 2.5) and the temperature is increasing. T&O concentrations quickly go down as the system thermally stratifies during the summer months of July through August and residence times significantly increase, moving plotted samples to the top right of Figure 2.5. September shows a great deal of variability attributable to large storm events and discharge into the reservoir, oscillating conditions back and forth along the x-axis. By October the second seasonal mixing of ECR is well underway and the plot shifts back to the very bottom, mimicking the state of the system in April but with increased nutrient concentrations. Based on this information and the timing of events illustrated by the time series (Figure 2.3) it seems that T&O concentrations are at their highest when the weather was more cold, windy, and rainy, and nutrient conditions on a decline following previously elevated nutrient concentrations by early-year discharges into ECR. T&O concentrations were at a minimum during a combination of high temperatures and low winds in the summer, then made a small resurgence under late season rains, colder temperatures, and little wind.

The microbial community at ECR is variable in composition, as different microbes thrive across depth and changing water column conditions throughout the season [Clercin and Druschel 2019]. Based on 16s rRNA gene sequencing (Clercin and Druschel, 2019) cyanobacteria are the dominant phylum in the bacterioplankton

community, constituting 36% of the bacteria present on average. Proteobacteria were a close second at 25%, but there are no known producers of GSM and MIB in the phylum. Actinobacteria averaged 7% of the community. Spearman cross-correlation analyses of 16s data and T&O concentrations tied GSM concentrations to genera from the Oscillatoriales and Nostocales orders of the cyanobacterial phylum, but not to any actinobacteria genera. Conversely, MIB was linked to genera from the Actinomycetales order of the actinobacteria phylum, but not any cyanobacterial genera (Clercin and Druschel, 2019). Cyanobacteria are known to outcompete other bloom microorganism communities at elevated water temperatures (Elliot *et al.* 2006, Jøhnk *et al.* 2008) and high temperatures generally improve their growth (Robarts and Zohary 1987). High temperatures can also assist in water column stratification where oxygen depletion in the hypolimnion stimulates nutrient release from the sediment-water interface (Søndergaard *et al.* 2003) which cyanobacteria are uniquely adapted to take advantage of through control of their buoyancy [Walsby *et al.* 2006]. Despite this, T&O concentrations at ECR are typically at their maximum during May when the waters are cold, well mixed, and highly turbid. Canonical correspondence analysis (CCA) showed that certain cyanobacteria such as *Planktothrix* and actinobacteria like *Streptomyces* did well under those water conditions and were therefore the most likely to be responsible for early season in situ production of GSM and MIB (Clercin and Druschel 2019).

The 2018 dataset is fundamentally different in how cell counts and/or biomass were measured compared to the work of Clercin and Druschel (2019). They determined community composition and abundance using metagenomics data, while here microscopy cell counts, Chl-*a*/PC, universal primer qPCR counts, and MIB/GSM synthase targeted

primer qPCR counts were used. Despite these large differences, nothing in the 2018 data appears to refute the conclusions of Clercin and Druschel (2019). Only weak links exist between any of the microscope cell counts and the T&O compounds across all analyses. Cyanophytes were correlated with GSM ($r_s = 0.21$, $p < 0.05$) and chlorophytes with MIB ($r_s = 0.18$, $p < 0.05$) but these relationships are small when compared to the relationships with qPCR counts or other measured parameters. This was to be expected and is unsurprising given that microscopy cannot enumerate actinobacteria without the use of genetic tagging techniques (Sekar *et al.* 2003). MIB (tied to actinobacteria) was the major T&O compound during 2018 and, consistent with Clercin and Druschel (2019), no significant links between measures of bloom biomass/density (Chl-*a*, PC, cell counts) and T&O concentrations have been found (Figures 2.4 and B2.4). Furthermore, cell counts are unimportant in the BRT model (Figure 2.6) and have relatively short vectors compared to environmental, weather, or nutrient parameters in the PCA and NMDS models (Figures B2.5 and 2.5) which supports their lack of a connection to T&O concentrations and overall reservoir conditions. Genetic data from qPCR analyses is a much better predictor of T&O production potential than bloom biomass or density, evidenced by cross-correlation analyses (Figures 2.4, B2.4) and the contribution of MIB qPCR counts to the BRT model (Figure 2.6 C). As discussed previously MIB formed a tighter correlation with GSM qPCR counts than GSM did, which could be the result of MIB and GSM co-occurrence or an artifact on GSM's low concentration and variability throughout 2018. Regardless, the PCA model (Figure B2.5) was able to group each T&O compound with its respective qPCR counts and separate the two groups, which supports

that the qPCR primers are functioning properly and indeed targeting separate organisms responsible for each compound.

Spatial variability of T&O compounds

The 2018 sampling regimen was designed in the hopes of capturing bloom hot spots and/or proliferation at ECR. Sampling was heavily focused on the northern basin (where Eagle Creek watershed drains into) with 12 locations, and sampling conducted on a bi-weekly to weekly basis. T&O concentrations, along with other measured variables, were visualized across space and time using pseudo-heat maps (heat-colored circles/dots overlaid on a satellite view of ECR) but no obvious patterns were visible (Figures B2.2 and B2.3). Further investigations were performed by including sampling location in the pairwise and multivariate analyses. Latitude and longitude were used in the creation of the PCA, NMDS, and BRT models (Figures B2.5, 2.5, 2.6) but these spatial coordinates as variables were found to be insignificant. Cross-correlation showed that latitude was loosely linked to Chl-*a* ($r_s = 0.31$, $p = 2.31x10^{-9}$) and several metals, but not significantly to GSM or MIB (Figure 2.4). This result suggests that bloom density tends to be greater the further north in ECR, but as shown with our data, algal biomass is not linked to T&O concentrations in this data nor has it been previously associated with T&O events [Clercin and Druschel 2019]. Additionally, there was no evidence of a connection between collected spectral data and the spatio-temporal variation of MIB and GSM. Reflectance values were only found to be loosely correlated with most parameters, the strongest correlations existing between Chl-*a* and several reflectance ratios ($r = 0.77$, $p < 0.05$). These ratios could serve in tracking the spatial variability of phytoplankton or cyanobacterial blooms but is unlikely to be useful in monitoring T&O concentrations.

Management Implications

The predilection for microorganisms such as *Planktothrix* and *Streptomyces* for cold, mixed, and highly turbid waters [Clercin and Druschel 2019] and the repeated findings of peak T&O concentrations at ECR when the water column matches these conditions, rule out common mitigation practices based on mixing and oxygenation. Attempts at mechanically mixing or oxygenating the water column could select for the organisms most responsible for high T&O concentrations, and ECR serves as a warning towards immediate implementation of these management strategies at other locations before the system in question is understood. Practices aimed at reducing eutrophication in the watershed are far more likely to succeed given the importance of the Eagle Creek watershed in introducing nutrients to the reservoir. Given the lack of detrimental outcomes, eutrophication reduction may serve as a first response tool in combating HABs and T&O events in other freshwater systems. Management solutions may also lie in manipulation of the reservoir level, and thus changing the residence time in the reservoir, by control of the outflow from the ECR dam. Reservoir water levels are frequently raised or lowered by several feet to assist in construction or maintenance projects. This could be harnessed to exert some control over HAB dynamics, either by flushing T&O compounds and blooms out of the reservoir directly or by making water column conditions unfavorable for phytoplankton growth. Rapid oscillations in the residence time of the reservoir, aided by manipulation of the water level, could drive enough variability in the water column conditions to prevent the establishment of microbes responsible for T&O production. This could prove an effective management solution for other freshwater reservoirs with similar dam controls as well.

References

Ahn, C.Y., S.H. Joung, S.K. Yoon, H.M. Oh. 2007. Alternative Alert System for Cyanobacterial Bloom, Using Phycocyanin as a Level Determinant. The Journal of Microbiology. 45: 98-104

Albarracín, V.H., A.L. Ávila, M.J. Amoroso, C.M. Abate. 2008. Copper removal ability by Streptomyces strains with dissimilar growth patterns and endowed with cupric reductase activity. FEMS Microbiol. Lett. 288: 141-148, doi: 10.1111/j.1574-6968.2008.01335.x

Andersen, I.M., T.J. Williamson, M.J. González, M.J. Vanni. 2020. Nitrate, ammonium, and phosphorus drive seasonal nutrient limitation of chlorophytes, cyanobacteria, and diatoms in a hyper-eutrophic reservoir. Limnol. Oceanogr. 65: 962-978, doi: 10.1002/lno.11363

Anderson, D. 2014. HABs in a changing world: a perspective on harmful algal blooms, their impacts, and research and management in a dynamic era of climactic and environmental change. Harmful Algae. 3-17.

Anderson, D.M., P.M. Gilbert, and J.M. Burkholder. 2002. Harmful Algal Blooms and Eutrophication: Nutrient Sources, Composition, and Consequences. Estuaries. 25: 704-726, doi: 140.182.176.10

APHA. 2000. Supplement to Standard Methods for the Examination of Water and Wastewater, 20th ed. APHA, AWWA, WPCF.

Arar, E.J. 1997. Method 446.0: In Vitro Determination of Chlorophylls a, b, c + c, and Pheopigments in Marine and Freshwater Algae by Visible Spectrophotometry. U.S. EPA. EPA/600/R-15/005.

Asquith, E.A., C.A. Evans, P.M. Geary, R.H. Dunstan, B. Cole. 2013. The role of Actinobacteria in taste and odour episodes involving Geosmin and 2-methyisoborneol in aquatic environments. Journal of Water Supply: Research and Technology-Aqua. 62(7): 452-467

Barker, T., H. MD. Irfanullah, and B. Moss. 2010. Micro-scale structure in the chemistry and biology of a shallow lake. Freshwater Biology. 55: 1145-1163, doi: 10.1111/j.1365-2427.2009.02339.x

Beisner, B.E., D.T. Haydon, K. Cuddington. 2003. Alternative stable states in ecology. Front. Ecol. Environ. 1: 376-382

Bentley, R., and R. Meganathan. 1981. GSM and Methylisoborneol Biosynthesis in Streptomycetes: Evidence for an Isoprenoid Pathway and its Absence in Non-differentiating Isolates. FEBS letters. 125: 220-22, doi: 10.1016/0014-5793(81)80723-5

Bishop, W.M., B.M. Johnson, J.H. Rodgers, Jr. 2014. Comparative responses of target and nontarget species to exposures of a copper-based algaecide. J. Aquat. Plant Manage. 52: 65-70

Cane, D.E., X. He, S. Kobayashi, S. Ömura, H. Ikeda. 2006. Geosmin biosynthesis in Streptomyces avermitilis. Molecular cloning, expression, and mechanistic study of the germacradienol/Geosmin synthase. The Journal of Antibiotics. 59(8): 471

Carmichael, W. W. 2001. Health Effects of Toxin-producing Cyanobacteria: 'The Cyano-HABs'. Human and Eco. Risk Assess. 7: 1393-1407, doi: 10.1080/20018091095087

Carpenter, S.R., N.F. Caraco, D.L. Correll, R.W. Howarth, A.N. Sharpley, V.H. Smith. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications. 8: 559-568

Chaffin, J.D., T.B. Bridgeman, D.L. Bade, C.N. Mobilian. 2014. Summer phytoplankton nutrient limitation in Maumee Bay of Lake Erie during high-flow and low-flow years. Journal of Great Lakes Research. 40: 524-531

Chattopadhyay, J., and R.R. Sarkar. 2002. A delay differential equation model on harmful algal blooms in the presence of toxic substances. IMA Journal of Mathematics Applied in Medicine and Biology. 19: 137-161

Christophoridis, C., K. Fytianos. 2006. Conditions Affecting the Release of Phoshorus from Surface Lake Sediments. J. Environ. Qual. 35: 1181-1192, doi: 10.2134/jeq2005.0213

Clercin, N.A. 2010. Monitoring Standard for Freshwater Algae: Analysis and Enumeration of Planktic Microalgae. CEES Protocol Standards.

Clercin, Nicolas André. 2019. Origin and Fate of Odorous Metabolites, 2-Methylisoborneol and Geosmin, in a Eutrophic Reservoir. Department of Earth Sciences, Indiana University

Clercin, N.A. and G.K. Druschel. 2019. Influence of Environmental Factors on Off-Flavor Metabolite Production by Bacteria in a Eutrophic Reservoir. Water Resources Research. 55: 5413-30, doi: 10.1029/2018WR023651

D'Angelo, E., J. Crutchfield, M. Vandiviere. 2001. Rapid, Sensitive, Microscale Determination of Phosphate in Water and Soil. J. Environ. Qual. 30: 2206-09, doi: 10.2134/jeq2001.2206

Davidson, K. and others. 2012. Harmful algal blooms: How strong is the evidence that nutrient ratios and forms influence their occurrence? Estuarine, Coastal, and Shelf Science. 115: 399-413, doi: 10.1016/j.ecss.2012.09.019

Delwiche, C.C. 1970. The Nitrogen Cycle. Scientific American. 223: 136-147

Dent, C.L., G.S. Cumming, S.R. Carpenter. 2002. Multiple states in river and lake ecosystems. Phil. Trans. R. Soc. Lond. 357: 635-645, doi: 10.1098/rstb.2001.0991

Diaz, R.J. 2001. Overview of Hypoxia around the World. Journal of Environmental Quality. 30: 275-281.

Dray, S. and J. Josse. 2015. Principal Component Analysis with Missing Values: a Comparative Survey of Methods. Plan Ecol. 216: 657–667, doi: 10.1007/s11258-014-0406-z

Elith, J., J.R. Leathwick, T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology. 77: 802-13, doi: 10.1111/j.1365-2656.2008.01390.x

Elliot, J.A., I.D. Jones, S.J. Thackeray. 2006. Testing the sensitivity of phytoplankton communities to changes in water temperature and nutrient load, in a temperate lake. Hydrobiologia. 559: 401-411

Fay, P. 1992. Oxygen Relations of Nitrogen Fixation in Cyanobacteria. Microbiological Reviews. 56: 340-73.

Flynn, K.J. 2010. Do external resource ratios matter?: Implications for modelling eutrophication events and controlling harmful algal blooms. Journal of Marine Systems. 83: 170-180.

Gerber, N., and H. Lechevalier. 1965. GSM, an Earthy-smelling Substance Isolated from Actinomycetes. Applied Microbio. 13: 935-38.

Giglio, S., J. Jiang, C.P. Saint, D.E. Cane, P.T. Monis. 2008. Isolation and characterization of the gene associated with Geosmin production in cyanobacteria. Environmental Science and Technology. 42(21): 8027-8032

Gitelson, A.A., J.F. Schalles, C.M. Hladik. 2007. Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. Remote Sensing of Environment. 109: 464-72, doi: 10.1016/j.rse.2007.01.016

Glibert, P.M., T.M. Kana, K. Brown. 2013. From limitation to excess: the consequences of substrate excess and stoichiometry for phytoplankton physiology, trophodynamics and biogeochemistry, and the implications for modeling. Journal of Marine Systems. 125: 14-28

Graham, J.L., G.M. Foster, T.J. Williams, M.D. Mahoney, M.R. May, K.A. Loftin. 2018. Water-quality conditions with an emphasis on cyanobacteria and associated toxins and taste-and-odor compounds in the Kansas River, Kansas, July 2012 through September 2016. U.S. Geological Survey Scientific Investigations Report. doi: 10.3133/sir20185089

Graham, J.L., K.A. Loftin, M.T. Meyer, A.C. Ziegler. 2010. Cyanotoxin Mixtures and Taste-and-Odor Compounds in Cyanobacterial Blooms from the Midwestern United States. Environ. Sci. Technol. 44: 7361-7368, doi: 10.1021/es1008938

Hamill, K. D. 2001. Toxicity in Benthic Freshwater Cyanobacteria (Blue-green Algae): First Observations in New Zealand. New Zealand Journal of Marine and Freshwater Research. 35: 1057-59, doi: 10.1080/00288330.2001.9517062

Harred, L.B. and L. Campbell. 2014. Predicting harmful algal blooms: a case study with *Dinophysis ovum* in the Gulf of Mexico. J. Plankton Res. 36: 1434-1445

Harris, T. D., V. H. Smith, J. L. Graham, D. B. Van de Waal, L. P. Tedesco, and N. Clercin. 2016. Combined Effects of Nitrogen to Phosphorus Ratios and Nitrogen Speciation on Cyanobacterial Metabolite Concentrations in Eutrophic Midwestern USA Reservoirs. Inland Waters. 6: 199-210, doi: 10.5268/IW-6.2.938

Heisler, J. and others. 2008. Eutrophication and Harmful Algal Blooms: A Scientific Consensus. Harmful Algae. 8: 3-13, doi: 10.1016/j.hal.2008.08.006

Horváth, H., A.W. Kovács, C. Riddick, M. Présing. 2013. Extraction Methods for Phycocyanin Determination in Freshwater Filamentous Cyanobacteria and Their Application in a Shallow Lake. Eur. J. Phycol. 48: 278-86, doi: 10.1080/09670262.2013.821525

Indiana Department of Environmental Management. 2018. Indiana Integrated Water Quality Monitoring and Assessment Report.

Jiang, J., X. He, D.E. Cane. 2007. Biosynthesis of the earthy odorant Geosmin by a bifunctional Streptomyces coelicolor enzyme. Nature Chemical Biology. 3: 711, doi: 10.1038/nchembio.2007.29

Jøhnk, K.D., J. Huisman, J. Sharples, B. Sommeijer, P.M. Visser, J.M. Stroom. 2008. Summer heatwaves promote blooms of harmful cyanobacteria. Global Change Biology. 14: 495-512, doi: 10.1111/j.1365-2486.2007.01510.x

Josse, J., and F. Husson. 2016. missMDA: A Package for Handling Missing Values in Multivariate Data Analysis. Journal of Statistical Software. 70.

Juttner, F. 1990. Monoterpenes and Microbial Metabolites in the Soil. Environ. Pollut. 68: 377-82.

Juttner, F., and S. B. Watson. 2007. Biochemical and Ecological Control of GSM and 2-methylisoborneol in Source Waters. Appl. Environ. Microbio. 73: 4395-406, doi: 10.1128/AEM.02250-06

Kansole, M. M.R. and T.-F. Lin. 2017. Impacts of Hydrogen Peroxide and Copper Sulfate on the Control of *Microcystis aeruginosa* and MC-LR and the Inhibition of MC-LR Degrading Bacterium *Bacillus* sp. Water. 9: 255, doi: 10.3390/w9040255

Koltsidou, Ioanna. 2019. Detection and Quantification of Taste and Odor Producing Bacteria in Eagle Creek Reservoir. Department of Biological Sciences, Purdue University.

Komatsu, M., M. Tsuda, S. Ōmura, H. Oikawa, and H. Ikeda. 2008. Identification and Functional Analysis of Genes Controlling Biosynthesis of 2-methylisoborneol. Proceedings of the National Academy of Sciences. 105: 7422-27, doi: 10.1073/pnas.0802312105

Kuzuyama, T. 2002. Mevalonate and nonmevalonate pathways for the biosynthesis of isoprene units. Bioscience, Technology, and Biochemistry. 66(8): 1619-1627

Lanciotti, E., C. Santini, E. Lupi, and D. Burrini. 2003. Actinomycetes, Cyanobacteria and Algae Causing Tastes and Odours in Water of the River Arno Used for the Water Supply of Florence. Journal of Water Supply: Research and Technology – Aqua. 52: 489-500, doi: 10.2166/aqua.2003.0044

Lean, D.R.S. and others. 1987. The Effects of Changes in Both the Abundance of Nitrogen and Phosphorus and Their Ratio on Lake Okaro Phytoplankton, with Comment on Six Other Central Volcanic Plateau Lakes. Marine and Freshwater Research. 21: 539-42, doi: 10.1080/00288330.1987.9516257

Leschine, S., K. Holwell, and E. Canale-Parola. 1988. Nitrogen Fixation by Anaerobic Cellulolytic Bacteria. Science. 242: 1157-59, doi: 10.1126/science.242.4882.1157

Li, L., D.L. Pascual, L.P. Tedesco, K.L. Randolph, R.E. Sengpiel, B.E. Hall, J.S. Wilson. 2006. Developing a Survey Tool for the Rapid Assessment of Blue-Green Algae in Central Indiana's Reservoirs. IDNR LARE Report, CEES Publication 2006-1.

Li, L., N. Wan, N. Gan, B. Xia, and L. Song. 2007. Annual Dynamics and Origins of the Odorous Compounds in the Pilot Experimental Area of Lake Dianchi, China. Water Sci. and Technol. 55: 43-50, doi: 10.2166/wst.2007.160

Ma, Z. M., Y. Niu, P. Xie, J. Chen, M. Tao, and X. W. Deng. 2013. Off-flavor Compounds from Decaying Cyanobacterial Blooms of Lake Taihu. J. Environ. Sci-China. 25: 495-501, doi: 10.1016/S1001-0742(12)60101-6

McKillup, S. and M.D. Dyar. 2010. Geostatistics Explained: an Introductory Guide for Earth Scientists, Cambridge University Press.

Miguéns, D., and E. Valério. 2015. The Impact of Some Microcystins on the Growth of Heterotrophic Bacteria from Portuguese Freshwater Reservoir. Limnetica. 34: 215-226, doi: 10.23818/limn.34.17

Ömür-Özbek, P., J.C. Little, A.M. Dietrich. 2007. Ability of humans to smell GSM, 2-MIB, and nonadienal in indoor air when using contaminated drinking water. Water Sci. and Tech. 55: 249-56, doi: 10.2166/wst.2007.186

Paerl, H. W., and J. Huisman. 2009. Climate Change: A Catalyst for Global Expansion of Harmful Cyanobacterial Blooms. Environ. Microbio. Reports. 1: 27-37, doi: 10.1111/j.1758-2229.2008.00004.x

Paerl, H. W., R. S. Fulton, P. H. Moisander, and J. Dyble. 2001. Harmful Freshwater Algal Blooms, with an Emphasis on Cyanobacteria. The Scientific World Journal. 1: 76-113, doi: 10.1100/tsw.2001.16

Pascual, D.L and L.P. Tedesco. 2004. Eagle Creek Reservoir: Responses to Algaecide Treatment. CEES Publication 2004-02.

Penfold, W.J. 1914. On The Nature Of Bacterial Lag. Journ. of Hyg. 14: 215-241

Perkins, R.G., and others. 2019. Managing taste and odour metabolite production in drinking water reservoirs: The importance of ammonium as a key nutrient trigger. Journal of Environmental Management. 244: 276-284, doi: 10.1016/j.jenvman.2019.04.123

Petterson, K. 1998. Mechanisms for internal loading of phosphorus in lakes. Hydrobiologia. 373/374: 21-25.

Petticrew, E.L., and J.M Arocena. 2001. Evaluation of iron-phosphate as a source of internal lake phosphorus loadings. The Science of the Total Environment. 266: 87-93

Pinel-Alloul, B. 1995. Spatial heterogeneity as a multiscale characteristic of zooplankton community. Hydrobiologia. 300/301: 17-42.

Pinel-Alloul, B., and A. Ghadouani. 2007. The Spatial Distribution of Microbes in the Environment, Springer.

Podduturi, R., N. Clausen, N.B. Petersen, N. O.G. Jørgensen, M.A. Petersen. 2018. Continuous collection of volatiles produced by *Streptomyces* grown on oatmeal agar by headspace extraction and GC-MS. doi: 10.3217/978-3-85125-593-5-100

Ramsing, N. B., H. Fossing, T. G. Ferdelman, F. Andersen, and B. Thamdrup. 1996. Distribution of Bacterial Populations in a Stratified Fjord (Mariager Fjord, Denmark) Quantified by In Situ Hybridization and Related to Chemical Gradients in the Water Column. Applied and Environ. Microbio. 62: 1391-1404.

Robarts, R.D., and T. Zohary. 1987. Temperature effects on photosynthetic capacity, respiration, and growth rates of bloom-forming cyanobacteria. New Zealand Journal of Marine and Freshwater Research. 21: 391-399, doi: 10.1080/00288330.1987.9516235

Rodriguez-Concepción, M. and A. Boronat. 2002. Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. Plant Physiology. 130(3): 1079-1089

Rolfe, M.D. and others. 2011. Lag Phase Is a Distinct Growth Phase That Prepares Bacteria for Exponential Growth and Involves Transient Metal Accumulation. Journal of Bacteriology. 194: 686-701

Sarada, R., M.G. Pillai, G.A. Ravishankar. 1999. Phycocyanin from Spirulina sp.: Influence of Processing of Biomass on Phycocyanin Yield, C-Phycocyanin Extraction from Spirulina platensis Wet Biomass Analysis of Efficacy of Extraction Methods and Stability Studies on Phycocyanin. Process Biochemistry. 34: 795.

Sarkar, R.R., B. Mukhopadhyay, R. Bhattacharyya, S. Banjerjee. 2007. Time lags can control algal bloom in two harmful phytoplankton-zooplankton system. Applied Mathematics and Computation. 186: 445-459, doi: 10.1016/j.amc.2006.07.113

Schultz, D. and R. Kishony. 2013. Optimization and control in bacterial Lag phase. BMC Biology. 11: 120

Sekar, R., A. Pernthaler, J. Pernthaler, F. Warnecke, T. Posch, R. Amann. 2003. An Improved Protocol for Quantification of Freshwater *Actinobacteria* by Fluorescence In Situ Hybridization. Applied and Environmental Microbiology. 69: 2928-2935

Shatwell, T., J. Koehler, and A. Nicklisch. 2008. Warming Promotes Cold-adapted Phytoplankton in Temperate Lakes and Opens a Loophole for Oscillatoriales in Spring. Global Change Biology. 14: 2194-2200, doi: 10.1111/j.1365-2486.2008.01630.x

Shen, X., H. Zhang, X. He, H. Shi, C. Stephan, H. Jiang, C. Wan, and T. Eicholz. 2019. Evaluating the treatment effectiveness of copper-based algaecides on toxic algae Microsystis aeruginosa using single cell-inductively coupled plasma-mass spectrometry. Analytical and Bioanalytical Chemistry. 411: 5531-5543.

Siegelman, H.W., J.H. Kycia, J.A. Hellebust, J.S. Craigie. 1978. Algal Biliproteins. Handbook of Phycological Methods: Physiological and Biochemical Methods. 72-78.

Šimek, K., M. Comerma, J.-C. García, J. Nedoma, R. Marcé, and J. Armengol. 2011. The Effect of River Water Circulation on the Distribution and Functioning of Reservoir Microbial Communities as Determined by a Relative Distance Approach. Ecosystems. 14: 1-14, doi: 10.1007/s10021-010-9388-4

Smil, V. 2000. Phosphorus in the Environment: Natural Flows and Human Interferences. Annu. Rev. Energy Environ. 25: 53-88.

Smith, L., M.C. Watzin, G. Druschel. 2011. Relating sediment phosphorus mobility to seasonal and diel redox fluctuations at the sediment-water interface in a eutrophic freshwater lake. Limnol. Oceanogr. 56: 2251-2264, doi: 10.4319/lo.2011.56.6.2251

Smith, V.H., G.D. Tilman, J.C. Nekola. 1999. Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. Environmental Pollution. 100: 179-196.

Søndergaard, M., B. Riike, B. Bullet, E. Jeppesen. 2013. Persistent internal phosphorus loading during summer in shallow eutrophic lakes. Hydrobiologia. 710: 95-107.

Søndergaard, M., J.P. Jensen, E. Jeppesen. 2003. Role of sediment and internal loading of phosphorus in shallow lakes. Hydrobiologia. 506-509: 135-145.

Srinivasan, R., and G. A. Sorial. 2011. Treatment of Taste and Odor Causing Compounds 2-methyl Isoborneol and GSM in Drinking Water: A Critical Review. J. Environ. Sci-China. 23: 1-13, doi: 10.1016/S1001-0742(10)60367-1

Stewart, W.D.P., F.R.S. T. Preston, H.G. Peterson, N. Christofi. 1982. Nitrogen cycling in eutrophic freshwaters. Phil. Trans. R. Soc. Lond. 296: 491-509

Su, M. and others. 2015. MIB-producing Cyanobacteria (Planktothrix sp.) in a Drinking Water Reservoir: Distribution and Odor Producing Potential. Water Research. 68: 444-53, doi: 10.1016/j.watres.2014.09.038

Tedesco, L.P., D.L. Pascual, L.K. Shrake, R.E. Hall, L.R. Casey, P.G.F. Vidon, F.V. Hernly, K.A. Salazar, R.C. Barr. 2005. Eagle Creek Watershed Management Plan: An Integrated Approach to Improved Water Quality. Eagle Creek Watershed Alliance, CEES Publication 2005-07, IUPUI, Indianapolis, 182p.

Tyc, O., and others. 2017. Exploring bacterial interspecific interactions for discovery of novel antimicrobial compounds. Microbial Biotechnology. 10: 910-925, doi: 10.1111/1751-7915.12735

Valiela, I. and others. 1992. Couplings of Watersheds and Coastal Waters: Sources and Consequences of Nutrient Enrichment in Waquoit Bay, Massachusetts. Estuaries. 15: 443-457

Vitousek, P.M and others. 1997. Human alteration of the global nitrogen cycle: sources and consequences. Ecological Applications. 7:737-750

Vörös, L., and J. Padisák. 1991. Phytoplankton biomass and chlorophyll-*a* in some shallow lakes in central Europe. Hydrobiologia. 215: 111-119

Walsby, A.E., F. Schanz, M. Schmid. 2005. The Burgundy-blood phenomenon: a model of buoyancy change explains autumnal waterblooms by *Planktothrix rubescens* in Lake Zürich. New Phytologist. 169: 109-122, doi: 10.1111/j.1469-8137.2005.01567.x

Wang, C.M. and D.E. Cane. 2008. Biochemistry and molecular genetics of the biosynthesis of the earthy odorant methylisoborneol in Streptomyces coelicolor. J. Am. Chem. Soc. 130(28): 8908-8909, doi: 10.1021/ja803639g

Watson, S.B. 2003. Cyanobacterial and eukaryotic alagal odour compounds: signals or by-products? A review of their biological activity. Phycologia. 42: 332-350

Watson, S. B., P. Monis, P. Baker, and S. Giglio. 2016. Biochemistry and Genetics of Taste-and odor-producing Cyanobacteria. Harmful Algae. 54: 112-27, doi: 10.1016/j.hal.2015.11.008

Wnorowski, A. 1992. Tastes and Odours in the Aquatic Environment: A Review. Water S. A. 18: 203-14.

Xu, H., H.W. Paerl, B. Qin, G. Zhu, G. Gao. 2010. Nitrogen and Phosphorus Inputs Control Phytoplankton Growth in Eutrophic Lake Taihu, China. Limnol. Oceanogr. 55: 420-32, doi: 10.4319/lo.2010.55.1.0420

Chapter 3 – Conclusions

We originally hypothesized that:

1. The spatial and temporal distributions of microbial cell, GSM, and MIB concentrations are controlled by hydrologic and chemical conditions within ECR.
2. Changes in nutrient levels and chemical conditions with the surface water control the formation and proliferation of HABs at ECR.

The goal of this study was to explore these two hypotheses using a spatially and temporally expansive dataset containing a suite of environmental parameters, with the idea that high density sampling would capture important information missing from past work at ECR. The culmination of all analyses suggests that water column conditions, algal activity, and T&O concentrations at ECR are largely determined by seasonal weather patterns and conditions within the watershed. MIB was by far the T&O compound of greatest concern during 2018, and the highest concentrations were associated with the low temperatures, high winds, high nutrient loads, and well-mixed water column typically seen between April and June at ECR. The primers developed to target the cyanobacteria and actinobacteria responsible for GSM and MIB, respectively, appear to detect genetic potential for T&O production, critical as cell counts and cellular densities appear unrelated to T&O events. T&O concentrations being the result of bloom death and decomposition was effectively ruled out, and strong evidence points to some form of N in the environment acting as a trigger for secondary metabolite production/release. A BRT model constructed from the data showed nitrate levels in the watershed as the most important predictor of MIB concentration, with low nitrate related to peak MIB. This falls in line with other studies which found low TN:TP ratios triggered

elevated T&O compounds in Midwestern reservoirs [Harris *et al.* 2016] and ECR [Clercin & Druschel 2019], though the inclusion of DN:TDP ratios in the 2018 dataset showed no relation. The importance of nutrient ratios should not be ruled out, however, as the DN:TDP ratio used is not fully consistent with typical TN:TP ratios and nitrate was the only form of N measured. Major shifts and high variability were observed in estimated residence times for ECR during the season, experiencing periods of river-like and lacustrine behavior corresponding to seasonal mixing and stratification, respectively.

We had hoped that visualizing nutrient, T&O, and cell concentrations across space would allow for the identification of "hotspots" within the reservoir or the tracking of bloom formation and proliferation. There was no evidence found of either phenomenon, though, and parameter differences between sampling sites is relatively small when compared to the degree of temporal variability. Spatial variability may indeed play a key role in T&O concentrations at ECR but cannot be captured at the scale of 2018's sampling regimen. Sampling sites were, at the very least, located over one hundred meters from one another. The northern most and southern most sites were separated by roughly 2 kilometers. At this scale it is highly probable that the dataset was more biased towards abiotic rather than biotic processes at ECR [Pinel-Alloul & Ghadouani 2007] and was better at describing general reservoir conditions than spatial differences. Furthermore, Clercin and Druschel (2019) demonstrated important spatial differences exist within the water column across depth, information which is absent if only surface waters are sampled. It may be that, at ECR, any important spatio-temporal variability driving T&O events is with depth and time rather than across the area of the reservoir's surface. If this is the case, then collecting and analyzing data via remote

sensing techniques will fail to capture important information and the monitoring of conditions can only truly be ascertained with physical sampling of ECR.

To further examine these conclusions, as well as assess the predictive capabilities of the BRT model, extensive sampling across a number of scales, including various spatial and temporal dimensions would be needed. A somewhat simplified sampling regimen, focused only on the parameters/variables identified as important in the BRT model but including a greater number of total samples across multiple years, would serve as an excellent independent dataset for verifying the model. More information about biotic processes as they relate to T&O compounds at ECR could be obtained by narrowing the spatial scale of sampling to bring individual sites closer to one another. Alternatively, sampling could be spread across more of the southern basin or include portions of the ECR watershed to further confirm our findings or bias the data towards large-scale abiotic processes.

Nonetheless, the outcomes of this study carry important ramifications for the future management of ECR. The strong influence that watershed discharge, nutrient load, and weather variables have in the multivariate (Figures 2.5 and B2.5) and BRT (Figure 2.6) models implies that only a handful of parameters are needed to describe the state of the system. While the current lack of a similarly sized independent dataset makes testing the predictive capabilities of the BRT model improbable, once these capabilities have been tested the basic parameters in the BRT model may form the basis for future monitoring and serve as a reference point for interventions. For example, delayed summer temperatures could push the cessation of the first mixing period further back into the year, leading to an extended peak T&O event. Likewise, the BRT model (Figure 2.6)

and time series (Figure 2.3) results show that decreases in nitrate within the watershed precede or co-occur with the largest T&O events, and the concentration at which this decrease triggers a T&O event may represent the threshold at which ECR shifts from P-limited to N-limited states. The findings by Harris *et al.* (2016) that T&O compounds were closely associated with low TN:TP ratios in Midwestern reservoirs, coupled with a similar observation by Clercin and Druschel (2019) at ECR, further support this idea. Close monitoring of nitrate levels at the watershed gage may serve as an early warning for future events. It should be noted that the analyses were performed again with an inclusion of DN:TDP ratios, but the results of cross-correlation and the multivariate models were unchanged. No correlations between measures of biomass (Chl-*a*/PC) or cell density (microscopy cell counts) and GSM/MIB were found in the dataset, and none of these parameters were significant contributors to the multivariate models. This further reinforces the idea that T&O production and concentrations at ECR are a response to changing environmental parameters rather than a side effect of high biomass/density. Groups like Pascual and Tedesco (2004) were interested in the possible contributions of the diatom community to T&O production, but the results of Clercin and Druschel (2019) as well as the results here do not indicate the involvement of diatoms on any level.

The predilection for microorganisms such as *Planktothrix* and *Streptomyces* for cold, mixed, and highly turbid waters [Clercin and Druschel 2019] and the repeated findings of peak T&O concentrations at ECR when the water column matches these conditions, rule out common mitigation practices based on mixing and oxygenation. Attempts at mechanically mixing or oxygenating the water column could select for the organisms most responsible for high T&O concentrations. Practices aimed at reducing

eutrophication in the watershed are far more likely to succeed given the importance of the Eagle Creek watershed in introducing nutrients to the reservoir. Management solutions may also lie in manipulation of the reservoir level by control of the outflow from the ECR dam. Reservoir water levels are frequently raised or lowered by several feet to assist in construction or maintenance projects. This could be harnessed to exert some control over HAB dynamics, either by flushing T&O compounds and blooms out of the reservoir directly or by making water column conditions unfavorable for phytoplankton growth. Rapid oscillations in the residence time of the reservoir, aided by manipulation of the water level, could drive enough variability in the water column conditions to prevent the establishment of microbes responsible for T&O production. Special considerations should be given to shoreline ecology at ECR if this approach is utilized, as the potential for unintentional habitat destruction does exist.

In summary, future management of the reservoir should focus on close monitoring of watershed conditions, particularly N loads, as an early detection and warning method to predict T&O events. Mechanical mixing or oxygenation of the water column at ECR should be avoided, as it may select for the most active T&O producers. Based on the importance of watershed discharge, water column conditions, and residence times at ECR, a more appropriate intervention may involve adjusting the water level of the reservoir. T&O compounds and algae could be flushed from the reservoir, or at the very least, the preferred niches of the T&O producers disrupted. We were unable to detect any significant spatial variability in T&O concentrations across the reservoir which could be an effect of sampling location/design. Later studies should take into consideration these results when designing a sampling/experiment regimen and measurements should be

taken of as many N and P species/forms as possible to address gaps in information of this current study.

References

Clercin, N.A. and G.K. Druschel. 2019. Influence of Environmental Factors on Off-Flavor Metabolite Production by Bacteria in a Eutrophic Reservoir. Water Resources Research. 55: 5413-30, doi: 10.1029/2018WR023651

Harris, T. D., V. H. Smith, J. L. Graham, D. B. Van de Waal, L. P. Tedesco, and N. Clercin. 2016. Combined Effects of Nitrogen to Phosphorus Ratios and Nitrogen Speciation on Cyanobacterial Metabolite Concentrations in Eutrophic Midwestern USA Reservoirs. Inland Waters. 6: 199-210, doi: 10.5268/IW-6.2.938

Pascual, D.L and L.P. Tedesco. 2004. Eagle Creek Reservoir: Responses to Algaecide Treatment. CEES Publication 2004-02

Pinel-Alloul, B., and A. Ghadouani. 2007. The Spatial Distribution of Microbes in the Environment, Springer

Appendix A

Supplemental figures for Chapter 1

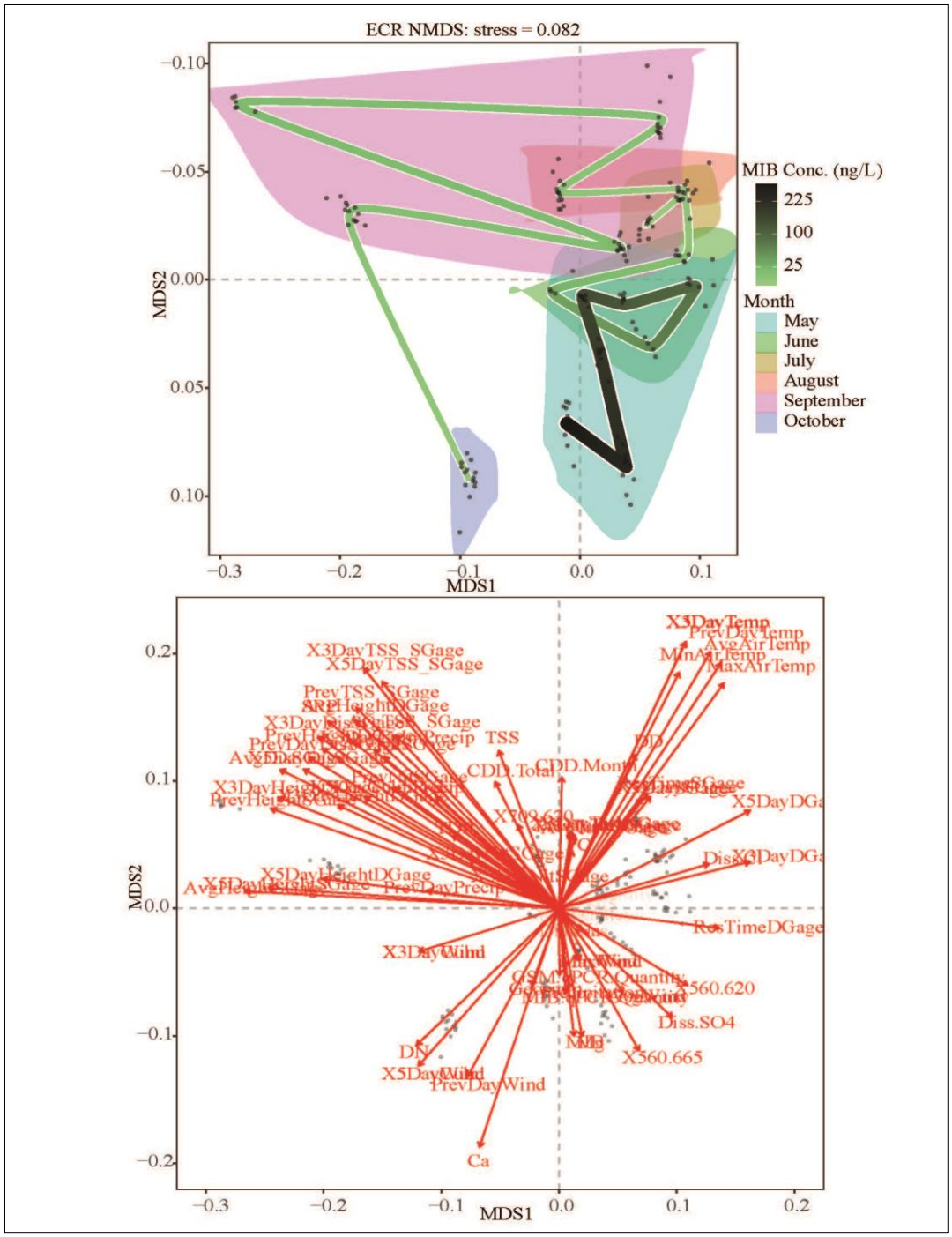

Figure A1.1: NMDS biplot (top) and vector space (bottom) for the trimmed (non-imputed) dataset.

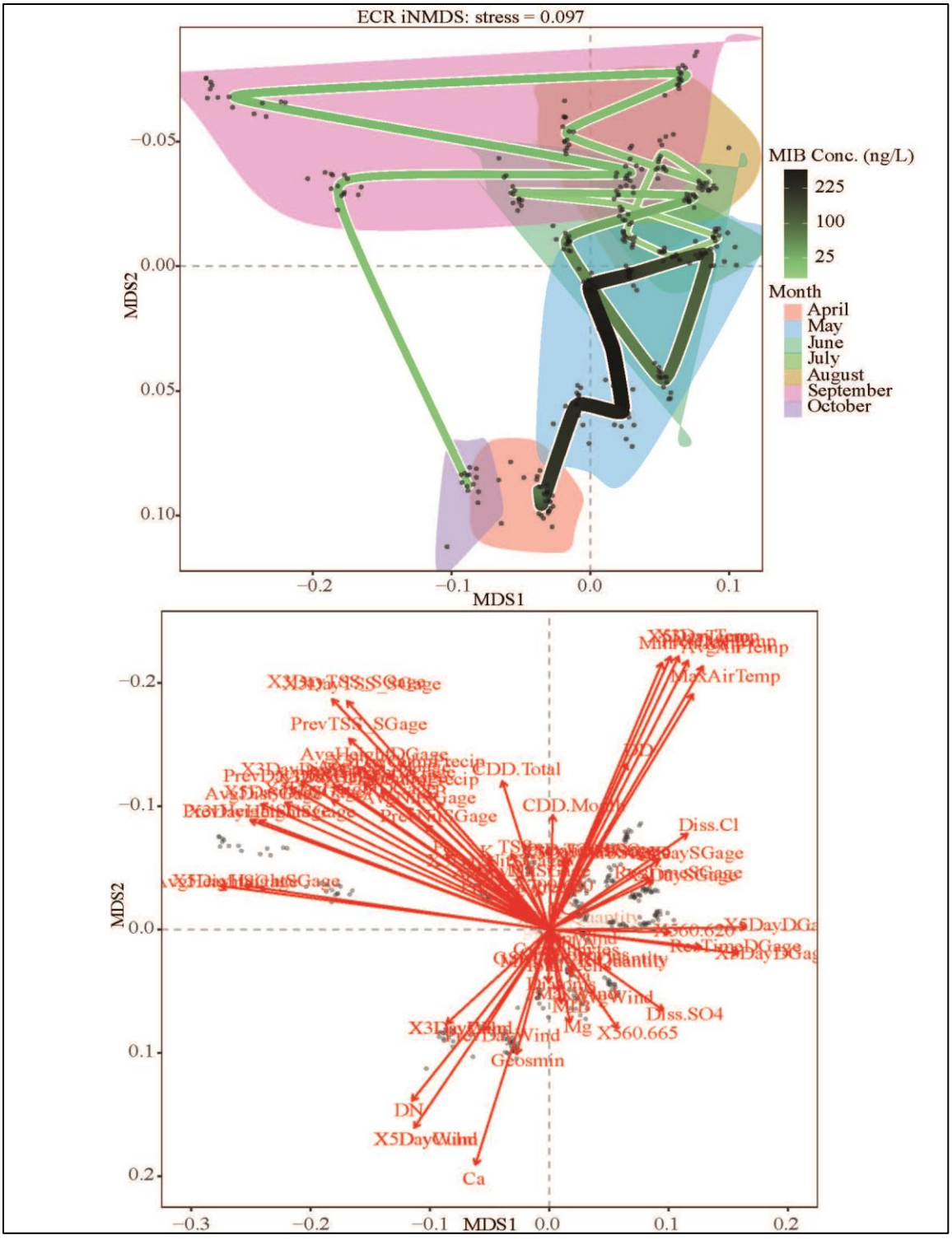

Figure A1.2: NMDS biplot (top) and vector space (bottom) for the imputed dataset.

Appendix B

Supplemental figures for Chapter 2

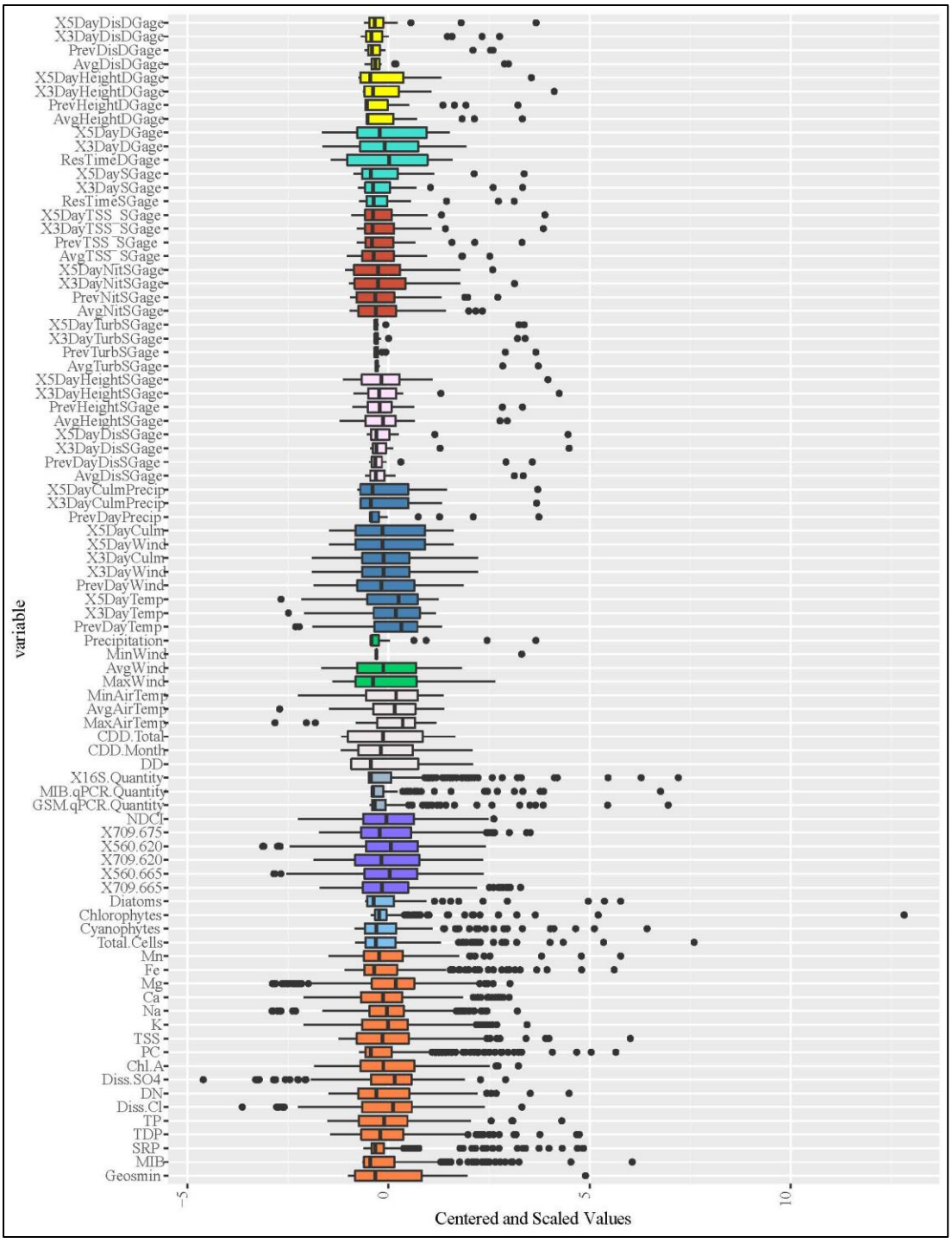

Figure B2.1: Box-and-whisker plots for a large subset of the variables, showing the distribution of data for each. Outliers are indicated by black dots on the graph.

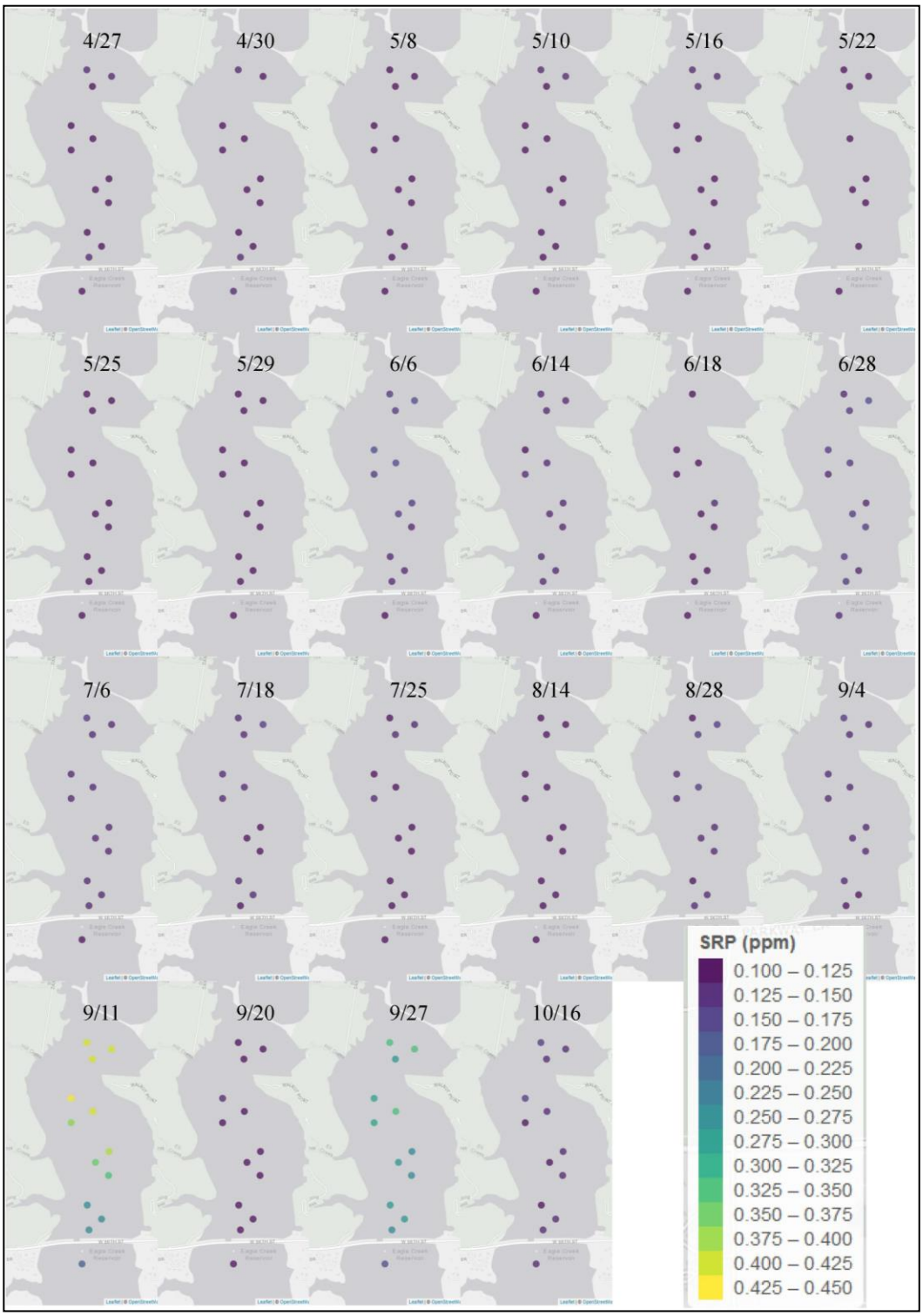

Figure B2.2: Bubble maps showing spatial and temporal variability in measured SRP concentration. Sampling points are indicated by bubbles, colored according to the scale indicated in the key.

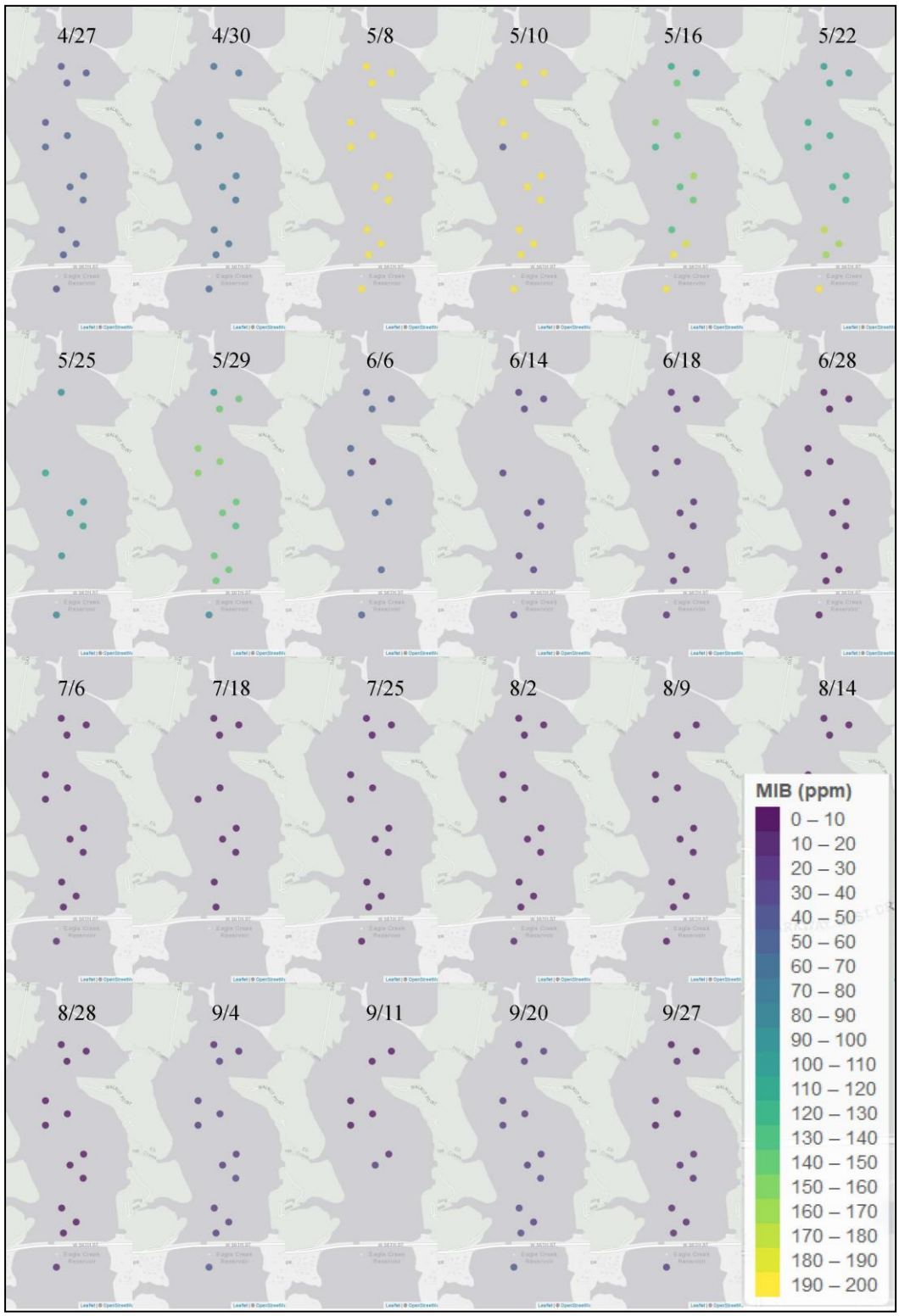

Figure B2.3: Bubble maps showing spatial and temporal variability in measured MIB concentration. Sampling points are indicated by bubbles, colored according to the scale indicated in the key.

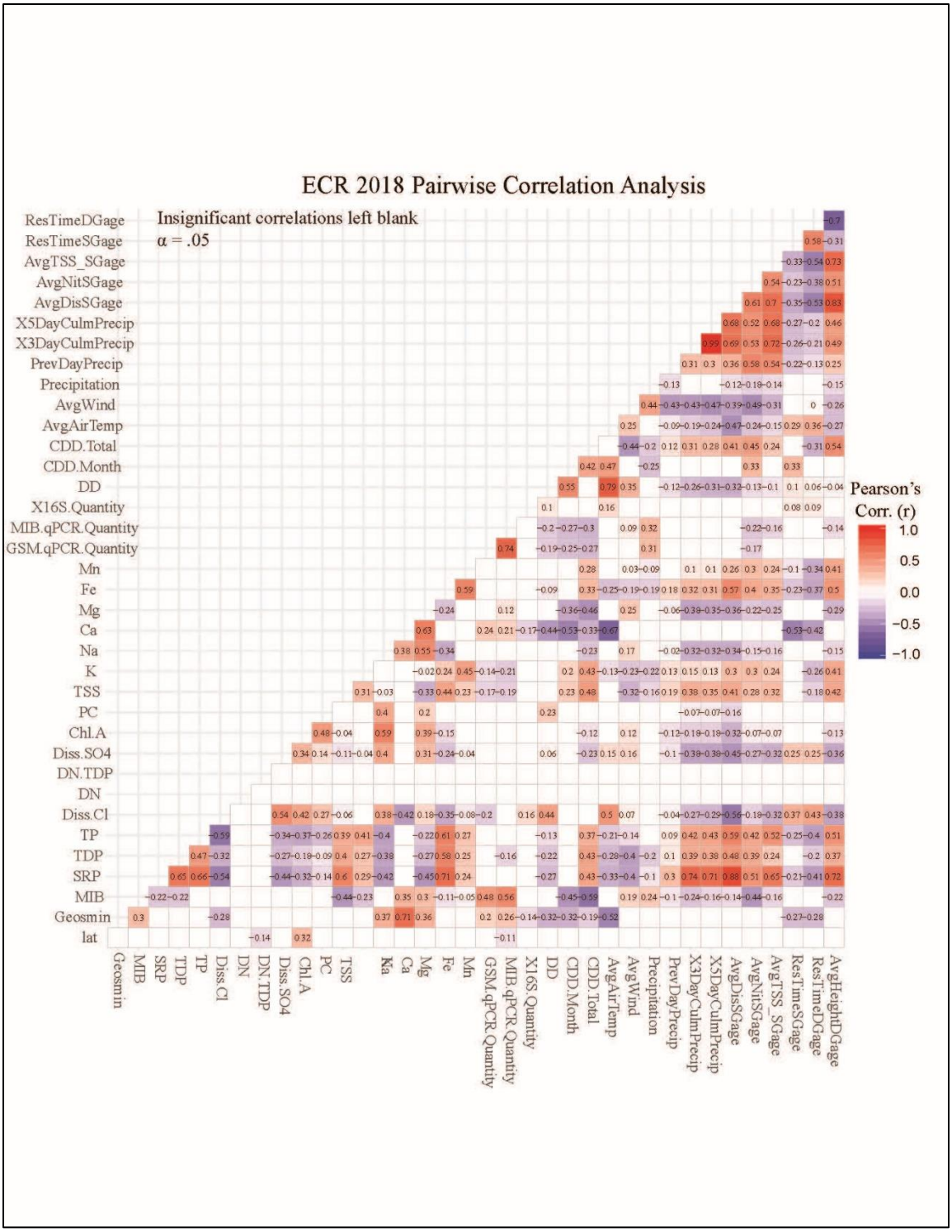

Figure B2.4: Pearson's Correlation matrix for a large selection of the dataset. Correlations with a p-value greater than 0.05 are left blank. The data and R script needed to re-create this and other matrices are available as part of the Appendix X and Y.

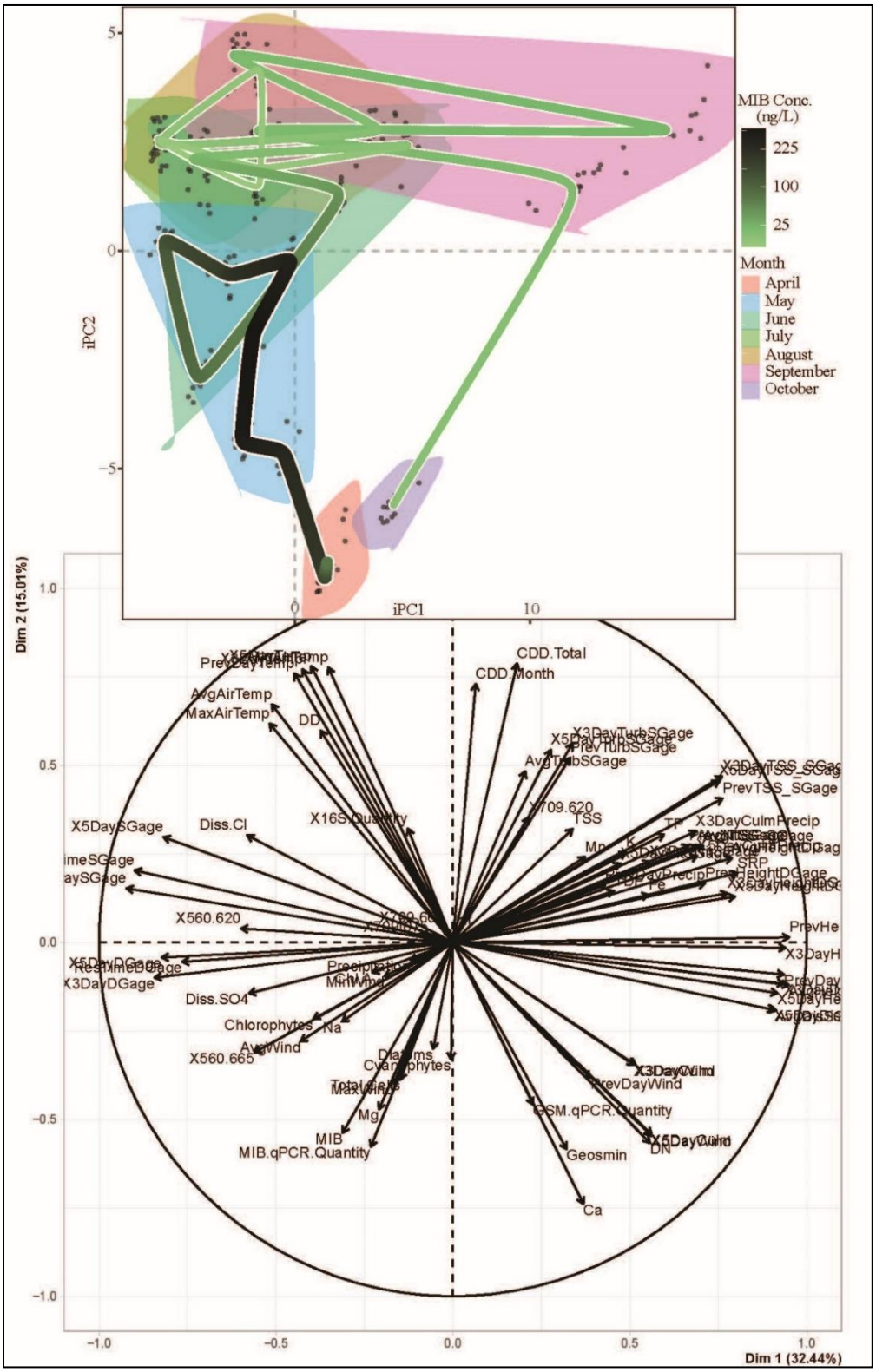

Figure B2.5: Results of PCA analysis on the imputed dataset. Biplot of samples (top) and vectors for the first two principal components (bottom) are shown.

Curriculum Vitae

Chase Steven Howard

Education

- Indiana University-Purdue University Indianapolis (IUPUI) (January 2018 – November 2020)
  Master of Science in Geology

- IUPUI (January 2013 – May 2017)
  Bachelor of Science in Chemistry, ACS certified
  Minors in Geology, Mathematics, and Physics

Research and teaching experience

- IUPUI (2018 - 2020)
  Graduate researcher
  Aqueous geochemistry research in the Earth Science department, focused primarily on nutrient cycling and algal bloom dynamics. Included field sampling, instrumental analysis, and experimental design. Managed and led undergraduate assistants in lab and field work.

- IUPUI (2018 - 2020)
  Teaching assistant
  Classes included: Earth Materials, Introduction to Geochemistry, and Applied Analytical Techniques in Geology. Graded assignments and reports, led lectures when professor was unavailable, assisted in lab assignments and activities, prepared materials for lab and class.

- IUPUI (2016 – 2017)
  Undergraduate researcher
  Cleaned and maintained a geochemistry laboratory, and kept track of chemical stock. Assisted in field sampling for several different projects and performed instrumental analysis.

Scholarships and awards

- Dean of Faculties Scholarship (2013 – 2016)
- Dean of Science Scholarship (2015 – 2016)
- School of Science Dean's List (2013 – 2016)
- IUPUI School of Science Scholars List Recipient (2013 – 2014)

Select presentations

- Howard, C.S., Kafantaris, F.C. A., & Druschel, G.K. Effect of bubbles and fluid flow on voltammetric signals and sulfur mineralization. Poster. *1st Geobiology Society Conference, Banff, June 2017.*
- Howard, C.S., Shukle, J., Kurek, K., Klatt, J., & Druschel, G.K. Sulfur speciation and gradients in mats from a dynamic sulfidic sinkhole, Middle Island, Lake Huron, MI. Poster. *5th Annual Midwest Geobiology Symposium, Cincinnati, October 2016.*
- Howard, C.S., Ogashawara, I., Koltsidou, I., & Druschel, G.K. Spatial and temporal variations of water quality in a eutrophic freshwater lake and their correlation to seasonal algal blooms. Poster. *7th Annual Midwest Geobiology Symposium, Evanston, October 2018.*
- Howard, C.S., & Druschel, G.K. Short-term temporal variations in the physiochemical conditions of a eutrophic freshwater lake. Poster. *8th Annual Midwest Geobiology Symposium, St. Louis, September 2019.*

Publications

- Ogashawara, I., Lin, L., Howard, C.S., & Druschel, G.K. (in prep). Monitoring phycocyanin with Landsat 8 / Operation Land Imager orange contra-band.
- Howard, C.S., Koltsidou, I., Druschel, G.K., & Picard, C.J. (in prep). Taste and odor event dynamics of a Midwestern freshwater reservoir.